

STIC-Biotech/ChemLib

95090

From: Rao, Manjunath N.
Sent: Tuesday, May 27, 2003 3:48 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/980771

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10D 01
Phone: 306-5681

Date: 5-27-03

Please search the following as soon as possible for application with serial number **09/980771**

SEQ ID NO: 2, 4, 6, 8 against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of all results.

SEQ ID NO:3, and amino acids 58-708 of SEQ ID NO:3, SEQ ID NO:5, 7, 9 against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Mail Box in 10D01
Crystal Mall 1, USPTO.

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 5/28
Date Completed: 6/10
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

This page Blank (uspto)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 08:28:25 ; Search time 3691.76 Seconds
(without alignments)
16743.883 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atgcgtgtgcctctaccag.....ccgccaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb_ba.*
- 2: gb_hlg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pin.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2124	100.0	2124	6	AX049323	Sequence
2	2124	100.0	3117	6	AX049322	Sequence
3	2122.4	99.9	3240	8	AF026420	Chlamydom
4	1953	91.9	1953	6	AX049325	Sequence
5	1593	75.0	1593	6	AX049329	Sequence
6	1314	61.9	1314	6	AX049327	Sequence
7	962.2	45.3	5856	8	AF433156	Chlamydom
8	537	25.3	2289	8	AF286320	Triticum
9	531.2	25.0	1812	8	AF486514	Hordeum v
10	531.2	25.0	2311	8	HWAXYR	Hordeum v
11	529.6	24.9	1827	8	AF486515	Barley mRNA
12	529.6	24.9	1827	8	AF486518	Hordeum v
13	527.4	24.8	2127	8	AF163319	Triticum
14	526.4	24.8	1812	8	AF486519	Hordeum v
15	526.4	24.8	1827	8	AF486516	Hordeum v
16	526.4	24.8	1827	8	AF486517	Hordeum v
17	525.8	24.8	1605	8	AF250137	Triticum
18	507.2	23.9	2028	8	AF113844	Triticum
19	507	23.9	2186	8	TAWAXYSS	Wheat waxy
20	507	23.9	2267	8	AY050174	Triticum
21	481.4	22.7	1801	8	AF113843	Triticum
22	478	22.5	2267	6	AR195560	Sequence
23	475.4	22.4	2115	8	AB089141	Setaria i
24	457.4	21.5	2287	8	SBU23945	Sorghum bic
25	449.2	21.1	1937	8	AB066093	Oryza sat
26	449.2	21.1	1937	8	AB066094	Oryza sat
27	447.6	21.1	2542	6	AR106491	Sequence
28	447.6	21.1	2542	8	OSWX	O. sativa wa
29	347	16.3	2168	8	MEGBSS	M. esculenta
30	344.8	16.2	2270	8	AF210699	Perilla f
31	332.6	15.7	2221	8	AB029546	Phaseolus
32	321.6	15.1	2211	8	AB071604	Ipomoea b
33	316.8	14.9	2161	6	AX394246	Sequence
34	313.8	14.8	3049	8	AF026421	Chlamydom
35	313.2	14.7	2345	8	AMA6293	Antirrhin
36	305.6	14.4	2124	8	IBU44126	Ipomoea bat
37	298	14.0	2092	8	AY094405	Arabidops
38	298	14.0	2115	8	AY088544	Arabidops
39	298	14.0	2192	8	AY123983	Arabidops
40	289.6	13.6	2107	8	AF097922	Astragalu
41	283.2	13.3	764	8	AF446083	Miscanthu
42	280.2	13.2	2035	8	PSSSTARSYN	P. sativum m
43	274.2	12.9	2081	8	AF109395	Triticum
44	274	12.9	764	8	AF446082	Microsteg
45	273.2	12.9	721	8	AF446081	Cleistach

ALIGNMENTS

RESULT 1	AX049323	AX049323	Sequence 2 from Patent WO0071734.	2124 bp	DNA	linear	PAT 12-JAN-2001
LOCUS	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
DEFINITION	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
ACCESSION	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
VERSION	AX049323.1	AX049323.1	GI:12226090				
KEYWORDS	AX049323.1	AX049323.1	GI:12226090				
SOURCE	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
ORGANISM	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
REFERENCE	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
AUTHORS	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
TITLE	AX049323	AX049323	Sequence 2 from Patent WO0071734.				
JOURNAL	AX049323	AX049323	Sequence 2 from Patent WO0071734.				

FEATURES
Location/Qualifiers
1..2124
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="fragment de la sequence compl te de l'DNAc codant
la GBSSI de Chlamydomonas reinhardtii"
1..>2124
CDS
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC21780.1"
/db_xref="GI:12226091"
/translation="MAVASTSPSSARPPIVINAASFGVKTKANQLRELARGSARKST
SRSAATGATGATCALDIVVAFAEAPWSKTGLGVDVTGGIPLTVKRGHRVMTIAPRY
DOYADWDVSVDVINGEKVRFHSIKKGVHRVWIDHPWFLAKVNGTKSKLYGPSG
ADLNHRKALFCAKAAEAAARVLPFGPDEDCVFNADHSAIPLVYLLEDEVPKQGF
TKASVLAHNLAFGRMWEAFKLPQAAFPDKLAFSDGYAKYVTEAPTEPMEDEKP
PLTGTYKKINLKGSIIAADKLVTSPNATEIAADAAGVDELDPVIRAKLEGIVN
MGDIEWNPDKTFLSAPYDQNSVAGKAAAEALQAEGLPVDPTAPLFAITGRLEE
OKGVDIILAAALPKILATPKVQILQILGTHAAVEKLVNAITGYKRGKGVKFSAPLIA
HMLTAGADPMLVPSFEPCGLIQLHAMHYGTVPVASTGLVDVTEGVTGPHMGALN
PKLLEADADALAATVRRASEVAGGRYPVMVANCISQDLSKSKPAQKNEGLLEEVY
GGGVATAKEELKVPVAEKIPEDLPVASTAPNTLAPVSAVSEGGAAAPKVGTTAPA
MGAWRATTPSGPSAAATPKVTYIKPALPATAKPKTAGIKLAGEAATTTSTSENGAASN
GNGNGASASKTSAAPLPLVSAATRKSA"
BASE COUNT 369 a 773 c 656 g 326 t
ORIGIN

Query Match 100.0%; Score 2124; DB 6; Length 2124;
Best local Similarity 100.0%; Pred. No. 7.1e-202;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGTTGCTTACACGCGCCCGCAGCAGCGCGGTCTCTATCGTCATCAATGCCGCG 60
DB 1 ATGGCTGTTGCTTACACGCGCCCGCAGCAGCGCGGTCTCTATCGTCATCAATGCCGCG 60
QY 61 TCGTTCGGTGTCAAGAGACCGGGAACAGCTGCTGGTGGTGGTTCGTTGGTTCGCGCA 120
DB 61 TCGTTCGGTGTCAAGAGACCGGGAACAGCTGCTGGTGGTGGTTCGTTGGTTCGCGCA 120
QY 121 CGCAAGTCCACCTCGCGCTCGGCTGTTACTGGTGCCACTGGTGCCTTCGCGCTGAC 180
DB 121 CGCAAGTCCACCTCGCGCTCGGCTGTTACTGGTGCCACTGGTGCCTTCGCGCTGAC 180
QY 181 ATCGTGATGGTGTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGCGCTGGCGATGTG 240
DB 181 ATCGTGATGGTGTGCTGAGGTGCGCCCTTGGTCCAAAGACGGCGCGCTGGCGATGTG 240
QY 241 ACTGGTGGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCTATGACCATTTGCCCT 300
DB 241 ACTGGTGGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCTATGACCATTTGCCCT 300
QY 301 CGCTACCAAGTACGCTGAGCTGGACACCTCGTGGTGGTGGTGGTGGTGGTGGTGG 360
DB 301 CGCTACCAAGTACGCTGAGCTGGACACCTCGTGGTGGTGGTGGTGGTGGTGGTGG 360
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGACCGCGTGTGGATTGACCAACC 420
DB 361 AAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGACCGCGTGTGGATTGACCAACC 420
QY 421 TGGTTCCTGGCCAAAGTCTGGGGCAAGACGGGTCCAAAGTGTACGGCCCGCTCCGGC 480
DB 421 TGGTTCCTGGCCAAAGTCTGGGGCAAGACGGGTCCAAAGTGTACGGCCCGCTCCGGC 480
QY 481 GCTGACTACCTGGACACCAAGCGCTTCGCCCTGTTCTCAAGGGCGCTATTGAGGCT 540
DB 481 GCTGACTACCTGGACACCAAGCGCTTCGCCCTGTTCTCAAGGGCGCTATTGAGGCT 540
QY 541 GCCCGGTGCTGCCCTTTCGGCCCGCGGAGACTGCGTTCCTTCGTTGGCCAAAGTGGG 600
DB 541 GCCCGGTGCTGCCCTTTCGGCCCGCGGAGACTGCGTTCCTTCGTTGGCCAAAGTGGG 600

QY 601 TCCGCCCTGTGTGCGCGCTCTGTGTAAGAGAGAGTACAGGCCCAAGGGCCAGTTCACCAAG 660
DB 601 TCCGCCCTGTGTGCGCGCTCTGTGTAAGAGAGAGTACAGGCCCAAGGGCCAGTTCACCAAG 660
QY 661 GCCAAGTCSGTGTGGGTATCCCAACATCGCTTCCAGGGCCGCGCATGTGGAGAGGCT 720
DB 661 GCCAAGTCSGTGTGGGTATCCCAACATCGCTTCCAGGGCCGCGCATGTGGAGAGGCT 720
QY 721 TTCAAGGACACGAAGCTGCGCCCGCGCTTTTACAAAGTGGCGCTTCTCGAGCGGCTAT 780
DB 721 TTCAAGGACACGAAGCTGCGCCCGCGCTTTTACAAAGTGGCGCTTCTCGAGCGGCTAT 780
QY 781 GCCAAGTCTTACACTGAGGCGCAACCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 GCCAAGTCTTACACTGAGGCGCAACCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 AAGACCTTACAGAGATCAACTGGCTGAAGGTGGCTATTATCGCGCGCGCAACAGCTGGT 900
DB 841 AAGACCTTACAGAGATCAACTGGCTGAAGGTGGCTATTATCGCGCGCGCAACAGCTGGT 900
QY 901 ACTGTGCGCCCAACTACGGGACCGAGATCGCTCCGATGCGCGCGCGGTGTGGAGCTG 960
DB 901 ACTGTGCGCCCAACTACGGGACCGAGATCGCTCCGATGCGCGCGCGGTGTGGAGCTG 960
QY 961 GACACCGTCTATCCCGCGCAAGGGCATTTGAGGGCATTTGAGGGCATTTGAGGGCAT 1020
DB 961 GACACCGTCTATCCCGCGCAAGGGCATTTGAGGGCATTTGAGGGCATTTGAGGGCAT 1020
QY 1021 TGGAAACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACAGAGAGAGAGAGAGAG 1080
DB 1021 TGGAAACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACAGAGAGAGAGAGAGAG 1080
QY 1081 GGCAAGCGCGCGCCCAAGAGAGGCGCTCGAGCGCGAGCTGGCGCTGCTGTGGAGAGAG 1140
DB 1081 GGCAAGCGCGCGCCCAAGAGAGGCGCTCGAGCGCGAGCTGGCGCTGCTGTGGAGAGAG 1140
QY 1141 GCGCCCTGTTTCGCGCTTATCGCGCGCGCTCGAGGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GCGCCCTGTTTCGCGCTTATCGCGCGCGCTCGAGGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GCGCGCTTGCACCAAGATCTGCGCCACCCCAAGGTGCAGATCGCATCTCTGGTACCGGC 1260
DB 1201 GCGCGCTTGCACCAAGATCTGCGCCACCCCAAGGTGCAGATCGCATCTCTGGTACCGGC 1260
QY 1261 AAGCGCGCTACGAGAGAGTGGTGAAGCCATCGGACCAAGTACAGAGGCGCGCGCAAG 1320
DB 1261 AAGCGCGCTACGAGAGAGTGGTGAAGCCATCGGACCAAGTACAGAGGCGCGCGCAAG 1320
QY 1321 GCGGTGTCAGTTCGCGCGCGCTGCGCACATGCTACCGCGCGCGCGAGTTCATG 1380
DB 1321 GCGGTGTCAGTTCGCGCGCGCTGCGCACATGCTACCGCGCGCGCGAGTTCATG 1380
QY 1381 CTGTTGCGCTTGCAGCGCTTGCAGCGCTGATCCAGCTGCAGCGCGCATCACTACGGT 1440
DB 1381 CTGTTGCGCTTGCAGCGCTTGCAGCGCTGATCCAGCTGCAGCGCGCATCACTACGGT 1440
QY 1441 ACCGTGCGGTGTGAGCGCTTCCACCGCGCGCTGCTGACACCGTCAAGAGGCGGTGAC 1500
DB 1441 ACCGTGCGGTGTGAGCGCTTCCACCGCGCGCTGCTGACACCGTCAAGAGGCGGTGAC 1500
QY 1501 GCGTTCACATGCGCGCGCTGAACCCGACAGGTGGAGAGTGCAGCGCGCGAGCGCGCTG 1560
DB 1501 GCGTTCACATGCGCGCGCTGAACCCGACAGGTGGAGAGTGCAGCGCGCGAGCGCGCTG 1560
QY 1561 GCCCGCACCGGTGCGCGGTGCCAGGAGTGTTCGCGCGCGCGCTACCCGAGAGTGTG 1620
DB 1561 GCCCGCACCGGTGCGCGGTGCCAGGAGTGTTCGCGCGCGCGCTACCCGAGAGTGTG 1620
QY 1621 GCCAACTGCATCAGGACGAGCTGCTCTGTTCCAAAGCGCGCGCGCAAGTGGAGAGGCTG 1680
DB 1621 GCCAACTGCATCAGGACGAGCTGCTCTGTTCCAAAGCGCGCGCGCAAGTGGAGAGGCTG 1680
QY 1681 CTGAGGAGAGTGGTGTACGGCAAGGGCGCGGTGGCGCGCGCGCGCGCGAGAGAGTCAAG 1740


```
Db 1681 CTGGAGGAGTGGTGTACGGCAAGGCGCGCTGGCCACCGCCAGAGAGAGATCAAG 1740
QY 1741 GTGCGCGTTCGCGAGAGATCCCGGCGACCTGCGCGCGTGTCTTACGCCGCCCAACACC 1800
Db 1741 GTGCGCGTTCGCGAGAGATCCCGGCGACCTGCGCGCGTGTCTTACGCCGCCCAACACC 1800
QY 1801 CTGAAGCCCGTGTCCGCGCTCGGTGAGGCGACGCGCGCGCGCGCGCGCGCGCGCGCG 1860
Db 1801 CTGAAGCCCGTGTCCGCGCTCGGTGAGGCGAACGCGCGCGCGCGCGCGCGCGCGCGCG 1860
QY 1861 ACCGCCCGCGCATGGGCGCGTGGCGCGCGACCAACCGCGCGCGCGCGCGCGCGCGCGCG 1920
Db 1861 ACCGCCCGCGCATGGGCGCGTGGCGCGCGACCAACCGCGCGCGCGCGCGCGCGCGCGCG 1920
QY 1921 GCCACCGCGAGGTGACCACTTACAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
Db 1921 GCCACCGCGAGGTGACCACTTACAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
QY 1981 GCTGGCGTCAAGTGGCGCGTGAAGCGTCCACCACTCGACCTCGAGAGAGCGCGCTGCC 2040
Db 1981 GCTGGCGTCAAGTGGCGCGTGAAGCGTCCACCACTCGACCTCGAGAGAGCGCGCTGCC 2040
QY 2041 TCCACGCGCAACGGCGCGTGGCGCGCGTCCCAAGACCTGGCGTCCCAAGCGCGCGTGGTC 2100
Db 2041 TCCACGCGCAACGGCGCGTGGCGCGCGTCCCAAGACCTGGCGTCCCAAGCGCGCGTGGTC 2100
QY 2101 TCCGCGCGCGCGCGAGTCCGCGC 2124
Db 2101 TCCGCGCGCGCGCGAGTCCGCGC 2124

RESULT 2
AX049322
LOCUS AX049322 3117 bp DNA linear PAT 12-JAN-2001
DEFINITION Sequence 1 from Patent WO0071734.
ACCESSION AX049322
VERSION AX049322.1 GI:12226089
KEYWORDS
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 3117)
AUTHORS D'Hulst,C. and Ball,S.
TITLE Starch granules containing a recombinant polypeptide of interest,
method for obtaining same and uses
JOURNAL Patent: WO 0071734-A 1 30-NOV-2000;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
source 1..3117
/organism="Chlamydomonas reinhardtii"
BASE COUNT 599 a 994 g 1011 g 513 t
ORIGIN
Query Match 100.0%; Score 2124; DB 6; Length 3117;
Best Local Similarity 100.0%; Pred. No. 6.3e-202;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGTGTTCCTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 15 ATGCGTGTTCCTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 74
QY 61 TCGTTCGGTGTCAAGAGACGCGGAACAGCTGCTGGTGAAGCTGCTGGTTCGCTCGCGCA 120
Db 75 TCGTTCGGTGTCAAGAGACGCGGAACAGCTGCTGGTGAAGCTGCTGGTTCGCTCGCGCA 134
QY 121 CGCAAGTCCACCTCGCGCTGGTGTACTGGTCCCACTGGTCCCACTGGCGCGCTGGAC 180
Db 135 CGCAAGTCCACCTCGCGCTGGTGTACTGGTCCCACTGGTCCCACTGGCGCGCTGGAC 194
QY 181 ATCGTGTATGTTGCTGTAGGTGCGCGCTTGGTCCCAAGAGCGGCGCGCTGGCGATGTG 240
```

```
Db 195 ATCGTGTATGTTGCTGTAGGTGCGCGCTTGGTCCCAAGAGCGGCGCGCTGGCGATGTG 254
QY 241 ACTGGTGGCTGCGTATTGAGCTGGTCAAGCGCGGCCACCGGCTCATGACCAATTGCCCT 300
Db 255 ACTGGTGGCTGCGTATTGAGCTGGTCAAGCGCGGCCACCGGCTCATGACCAATTGCCCT 314
QY 301 CGCTAGCACAGTACGCTGACGCTGGGACACCTCGGTGGTGGTGACATCATGGGCGAG 360
Db 315 CGCTAGCACAGTACGCTGACGCTGGGACACCTCGGTGGTGGTGACATCATGGGCGAG 374
QY 361 AAGTTCGGTACTTCCACTCCATCAAGAGGCGTGCACCGCTTACGGCCCCCGCTCCGCG 420
Db 375 AAGTTCGGTACTTCCACTCCATCAAGAGGCGTGCACCGCTTACGGCCCCCGCTCCGCG 434
QY 421 TGGTTCCTGGCCCAAGGCTGGGCGAGACCGGCTCCCAAGCTGACGGCCCCCGCTCCGCG 480
Db 435 TGGTTCCTGGCCCAAGGCTGGGCGAGACCGGCTCCCAAGCTGACGGCCCCCGCTCCGCG 494
QY 481 GCTGACTACTGGCAACCAACGCGCTTCGCGCTTCTGCAAGGCGCGCTATTGAGGCT 540
Db 495 GCTGACTACTGGCAACCAACGCGCTTCGCGCTTCTGCAAGGCGCGCTATTGAGGCT 554
QY 541 GCCCGGCTGCTGCCCTTCGGCCCCGGGAGGACTGCGTCTTCTGCGCCCAACGACTGGCAC 600
Db 555 GCCCGGCTGCTGCCCTTCGGCCCCGGGAGGACTGCGTCTTCTGCGCCCAACGACTGGCAC 614
QY 601 TCCGCGCTGGTGGCGCGTCTGCTGAAGACGAGTACGAGCCCAAGGGCGAGTTCACCAAG 660
Db 615 TCCGCGCTGGTGGCGCGTCTGCTGAAGACGAGTACGAGCCCAAGGGCGAGTTCACCAAG 674
QY 661 GCCAAGTGGTGGCTTATCCACAACATCGCTTCCAGGGCGCGATGTTGGGAGGAGCT 720
Db 675 GCCAAGTGGTGGCTTATCCACAACATCGCTTCCAGGGCGCGATGTTGGGAGGAGCT 734
QY 721 TTCAAGACACGAGCTGCCCGAGCGCGCTTTCGAAGCTGGCGCTTTCGAGACGGCTAT 780
Db 735 TTCAAGACACGAGCTGCCCGAGCGCGCTTTCGAAGCTGGCGCTTTCGAGACGGCTAT 794
QY 781 GCCAAGTTCACGTGAGGCGCACCCCGCTGAGGAGGAGGAGAGAGCCCGCTGACGGGA 840
Db 795 GCCAAGTTCACGTGAGGCGCACCCCGCTGAGGAGGAGGAGAGAGCCCGCTGACGGGA 854
QY 841 AAGACTACAAGAGATCAACTGGCTGAAGGTGCATTATCGCGCGCGCAAGTGGTG 900
Db 855 AAGACTACAAGAGATCAACTGGCTGAAGGTGCATTATCGCGCGCGCAAGTGGTG 914
QY 901 ACTGTGCGCCCACTACGCGACGAGATCGCTGCCGATGCGCGCGGCGGTGGAGCTG 960
Db 915 ACTGTGCGCCCACTACGCGACGAGATCGCTGCCGATGCGCGCGGCGGTGGAGCTG 974
QY 961 GACACCGTCAATCGCGCGCAAGGCGATTGAGGCGATTGAAAGCGCATGAGATTGAGGAG 1020
Db 975 GACACCGTCAATCGCGCGCAAGGCGATTGAGGCGATTGAAAGCGCATGAGATTGAGGAG 1034
QY 1021 TGAACCCCAAGACCGCAAGTTCCTGTCTCGCGCTTACGACCAAGAGAGAGTTCACGCC 1080
Db 1035 TGAACCCCAAGACCGCAAGTTCCTGTCTCGCGCTTACGACCAAGAGAGAGTTCACGCC 1094
QY 1081 GGCAGGCGCGCGCGCAAGGAGGCGCTTGCAGCGCGAGCTGGCGCTGCTGTGGACCCCAAC 1140
Db 1095 GGCAGGCGCGCGCGCAAGGAGGCGCTTGCAGCGCGAGTGGCGCTGCTGTGGACCCCAAC 1154
QY 1141 GCCCGCGTGTGGCTTTCATCGCGCGCTTGCAGGAGGAGGAGTGGGACATCATCTG 1200
Db 1155 GCCCGCGTGTGGCTTTCATCGCGCGCTTGCAGGAGGAGGAGTGGGACATCATCTG 1214
QY 1201 GCGCGCTGCCAAGATCTTGGCGACCCCGCAAGGTGCAGATCGCCATCTCTGGGTACCGCG 1260
Db 1215 GCGCGCTGCCAAGATCTTGGCGACCCCGCAAGGTGCAGATCGCCATCTCTGGGTACCGCG 1274
QY 1261 AAGCGCGCTACGAGAGCTGGTGAACGCGATCGGCAACAGTACAAAGGCGCGCGCAAG 1320
```


LOCUS AX049325 1953 bp DNA linear PAT 12-JAN-2001
DEFINITION Sequence 4 from Patent WO0071734.
ACCESSION AX049325
VERSION AX049325.1 GI:12226092
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1953)
AUTHORS D'Hulst,C. and Ball,S.
TITLE Starch granules containing a recombinant polypeptide of interest,
method for obtaining same and uses
JOURNAL Patent: WO 0071734-A 4 30-NOV-2000;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
Location/Qualifiers
1..1953
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="fragment de l'ADN complet codant la GHSSI de
Chlamydomonas reinhardtii et codant la proteine GBSSI
mature."
CDS
1..>1953
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="GI:12226093"
/db_xref="GI:12226093"

Query Match 91.9%; Score 1953; DB 6; Length 1953;
Best Local Similarity 100.0%; Pred. No. 6.1e-185;
Matches 1953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 GCCTGGACATCGTGTGCTGCTGAGTGGCCCTTGGTCCAGACGGCGGCTG 231
DB 1 GCCTGGACATCGTGTGCTGCTGAGTGGCCCTTGGTCCAGACGGCGGCTG 60
QY 232 GCGATGTGACTGGTGGCTGCTGCTATTGAGTGGTCAAGCGCGCCACCGCGTCATGACC 291
DB 61 GCGATGTGACTGGTGGCTGCTGCTATTGAGTGGTCAAGCGCGCCACCGCGTCATGACC 120
QY 292 ATTGCCCCCTGCTACGACGAGTACGCTGAGCCCTGGGACACCTCGGTGGTGGACATC 351
DB 121 ATTGCCCCCTGCTACGACGAGTACGCTGAGCCCTGGGACACCTCGGTGGTGGACATC 180
QY 352 ATGGGCGAGAAGTCCGCTACTTCCATCCATCAAGAAGGGCGTGCACCGCGTGTGGATT 411
DB 181 ATGGGCGAGAAGTCCGCTACTTCCATCCATCAAGAAGGGCGTGCACCGCGTGTGGATT 240
QY 412 GACACCCCTGGTCTCTGGCCAAAGTCTGGGGCAAGACCGCGTCCAAAGTGTACGGCCCC 471
DB 241 GACACCCCTGGTCTCTGGCCAAAGTCTGGGGCAAGACCGCGTCCAAAGTGTACGGCCCC 300
QY 472 CGCTCCGGGCTGACTACCTGGACACCAAGCGCTTCCGCTGTTCTCGAAGGGCGCT 531
DB 301 CGCTCCGGGCTGACTACCTGGACACCAAGCGCTTCCGCTGTTCTCGAAGGGCGCT 360
QY 532 ATTGAGGCTCCCGCGTGTGCTGCTGGCCCGGCGGAGGACTGCGTCTTCGTGGGCAAC 591
DB 361 ATTGAGGCTCCCGCGTGTGCTGCTGGCCCGGCGGAGGACTGCGTCTTCGTGGGCAAC 420

QY 592 GACTGGCACTCCGCCCTGGTGGCCGCTCTGCTCAAGGACGAGTACCAAGCCCAAGGGCCAG 651
DB 421 GACTGGCACTCCGCCCTGGTGGCCGCTCTGCTCAAGGACGAGTACCAAGCCCAAGGGCCAG 480
QY 652 TTCACCAAGGCCAAGTGGTGGCTATCCACAACATCGCTTCCAGGGCCGATGG 711
DB 481 TTCACCAAGGCCAAGTGGTGGCTATCCACAACATCGCTTCCAGGGCCGATGG 540
QY 712 GAGGAGCTTTCAGGACACGAGCTGCCAGGCGCCCTTGCACAGCTGGCTTCTCG 771
DB 541 GAGGAGCTTTCAGGACACGAGCTGCCAGGCGCCCTTGCACAGCTGGCTTCTCG 600
QY 772 GAGGAGCTTTCAGGAGTTTACACTGAGGCGCACCCCATGAGGAGGACGAGAGCCCGG 831
DB 601 GAGGAGCTTTCAGGAGTTTACACTGAGGCGCACCCCATGAGGAGGACGAGAGCCCGG 660
QY 832 CTGACGGGAAGACCTTACAAGAGATCACTGGCTGAAGGTTGGCATTTATCGCGCCGAC 891
DB 661 CTGACGGGAAGACCTTACAAGAGATCACTGGCTGAAGGTTGGCATTTATCGCGCCGAC 720
QY 892 AAGTGTGTGACTGTGTGCGCCCACTACGCGACCGAGATCGTCCGATGCGCGCGGGT 951
DB 721 AAGTGTGTGACTGTGTGCGCCCACTACGCGACCGAGATCGTCCGATGCGCGCGGGT 780
QY 952 GTGAGCTGGACACCGCTCATCCGCGCCCAAGGGCATTTGAGGSCATTTGTAAAGGATGGAC 1011
DB 781 GTGAGCTGGACACCGCTCATCCGCGCCCAAGGGCATTTGAGGSCATTTGTAAAGGATGGAC 840
QY 1012 ATTGAGGAGTGGAAACCCCAAGACCGGACAAAGTTCCTGCTGGGCCCTACGACCAAGACG 1071
DB 841 ATTGAGGAGTGGAAACCCCAAGACCGGACAAAGTTCCTGCTGGGCCCTACGACCAAGACG 900
QY 1072 GTCTACCGCGGCAAGGCGCGCCGCAAGGAGGCGCTGAGGCGGAGTGGCTGCTGCTG 1131
DB 901 GTCTACCGCGGCAAGGCGCGCCGCAAGGAGGCGCTGAGGCGGAGTGGCTGCTGCTG 960
QY 1132 GACCCCAACCGCCCTGTTGCTTTCATCGCGCGCTTGGAGGAGCAAGAGGTGTGGAC 1191
DB 961 GACCCCAACCGCCCTGTTGCTTTCATCGCGCGCTTGGAGGAGCAAGAGGTGTGGAC 1020
QY 1192 ATCATCTGGCGCGCTTGGCCCAAGATCTTGGCCACACCCCAAGGTGGACATCGCATCTG 1251
DB 1021 ATCATCTGGCGCGCTTGGCCCAAGATCTTGGCCACACCCCAAGGTGGACATCGCATCTG 1080
QY 1252 GGTACCGGCAAGGCGCGCTTACGAGAGTGTGTAAAGCCCATCGGACCAAGTCAAGAGGC 1311
DB 1081 GGTACCGGCAAGGCGCGCTTACGAGAGTGTGTAAAGCCCATCGGACCAAGTCAAGAGGC 1140
QY 1312 CGGCGCAAGGCGGTGTGCTTCTCGGCGCCCTGCGGCGCACATGCTCACCGCGCGCC 1371
DB 1141 CGGCGCAAGGCGGTGTGCTTCTCGGCGCCCTGCGGCGCACATGCTCACCGCGCGCC 1200
QY 1372 GACTTATGTGTGTGCTTCTCGGCGCTTACGAGCCCTGCGGCGCTGATCCAGCTGCAGCCCATG 1431
DB 1201 GACTTATGTGTGTGCTTCTCGGCGCTTACGAGCCCTGCGGCGCTGATCCAGCTGCAGCCCATG 1260
QY 1432 CACTTACGCTACCGTGGCGCTGTAGCTTCCACCGCGCGCTGGTTCGACACCGTCAAGAG 1491
DB 1261 CACTTACGCTACCGTGGCGCTGTAGCTTCCACCGCGCGCTGGTTCGACACCGTCAAGAG 1320
QY 1492 GCGGTACCGGCTTCCACATGGGCGCCCTGNAACCCCGACAGCTGGAGGCTGAGCGCC 1551
DB 1321 GCGGTACCGGCTTCCACATGGGCGCCCTGNAACCCCGACAGCTGGAGGCTGAGCGCC 1380
QY 1552 GACGCGCTGCGCGCCACCGTGGCGCTTCCAGCGAGGTGTTTGGCGGCGCGCTACGCC 1611
DB 1381 GACGCGCTGCGCGCCACCGTGGCGCTTCCAGCGAGGTGTTTGGCGGCGCGCTACGCC 1440
QY 1612 GAGATGTGGCCAACTGCATCAGCAGAGCTGCTGCTGGTCCAAAGCCCGCCGCAAGTGG 1671
DB 1441 GAGATGTGGCCAACTGCATCAGCAGAGCTGCTGCTGGTCCAAAGCCCGCCGCAAGTGG 1500


```

Db 901 GTCTACGGCGGCAAGCGCGCCCAAGAGGAGGCGCTCGAGCGCGAGCTGGCGCTGCTGTG 960
QY 1132 GACCCACCGCCCCCTGTTGCGCTTCATCGCGCGCCCTGGAGGACAGAAAGGCTGTGAC 1191
Db 961 GACCCACCGCCCCCTGTTGCGCTTCATCGCGCGCCCTGGAGGACAGAAAGGCTGTGAC 1020
QY 1192 ATCAPCTCGCGCGCGCTGCCCAAGATCTCGGCCACCCCAAGGTCAGATGCCATCCTG 1251
Db 1021 ATCATCTGCGCGCGCTGCCCAAGATCTCGGCCACCCCAAGGTCAGATGCCATCCTG 1080
QY 1252 GGTACCGCAAGGGCGGCTACGAGAAGCTGTGACGCGCATCGGCACCAAGTACAAAGGC 1311
Db 1081 GGTACCGCAAGGGCGGCTACGAGAAGCTGTGACGCGCATCGGCACCAAGTACAAAGGC 1140
QY 1312 CGCGCAAGGGCGGCTCAAGTCTCGCGCGCCCTGGCGGCACATGCTCACCGCGCGGCC 1371
Db 1141 CGCGCAAGGGCGGCTCAAGTCTCGCGCGCCCTGGCGGCACATGCTCACCGCGCGGCC 1200
QY 1372 GACTTCATGCTGGTGGCGCTCGCGCTTCGAGCGCTCGCGCGCTGATCCAGCTGCACGCCATG 1431
Db 1201 GACTTCATGCTGGTGGCGCTCGCGCTTCGAGCGCTCGCGCGCTGATCCAGCTGCACGCCATG 1260
QY 1432 CACTACGCTACCGTGGCGCTGGTACGCTCCACCGCGCGCTGGTGCACACCGTC 1485
Db 1261 CACTACGCTACCGTGGCGCTGGTACGCTCCACCGCGCGCTGGTGCACACCGTC 1314

RESULT 7
AF433156 5856 bp DNA linear PLN 23-AUG-2002
LOCUS Chlamydomonas reinhardtii granule-bound starch synthase I (ST2)
DEFINITION gene, complete cds.
ACCESSION AF433156
VERSION AF433156.1 GI:16755882
KEYWORDS Chlamydomonas reinhardtii.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 5856)
AUTHORS Watebled,F., Buleon,A., Bouchet,B., Rai,J.P., Lienard,L.,
Delvalle,D., Binderup,K., Dauvillee,D., Ball,S. and D'Hulst,C.
TITLE Granule-bound starch synthase I: a major enzyme involved in the
biogenesis of B-crystallites in starch granules
JOURNAL Eur. J. Biochem. 269 (15), 3810-3820 (2002)
PUBMED 12153578
REFERENCE 2 (bases 1 to 5856)
AUTHORS Watebled,F., Ball,S.G. and D'Hulst,C.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2001) UMR8576 du CNRS, Laboratoire de Chimie
Biologique, Universite des Sciences et Technologies de Lille,
Villeneuve d'Ascq CEDEX 59655, France
FEATURES
source location/Qualifiers
1..5856
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
1343..>4911
/gene="ST2"
1343..>4911
/gene="ST2"
Join(1343..1781,2151..2340,2484..2618,2815..2973,
3186..3373,3582..4117,4295..>4911)
/gene="ST2"
/product="granule-bound starch synthase I"
join(1480..1781,2151..2340,2484..2618,2815..2973,
3186..3373,3582..4117,4295..4911)
/gene="ST2"
/note="similar to the sequence deposited in GenBank
Accession Number AF026420; GBSSI"
/codon_start=1
/product="granule-bound starch synthase I"
/protein_id="AAL28128.1"
/db_xref="GI:16755883"
/translation="MAVASTSRPSSARPPIVNAASFGVKKTKANLLRELARGSARKST
SRSAVTGATGATCALDIVMVAEAVAPWSKTLGLDVTGGLPIELVKGHRHVMIIAPRY

```

```

DOYADAWDTSVVIMGEKRYFHSIKKKVHRWIDHPWFLAKVWMTGSKLYGPRSG
ADYLDNHRKFAFCFAAEEARVLPFGEGDCEVFVANDHWSALVPVLKLDYQPKQF
PLTKSVLAIHNIATGGRWEEAFKDTKLPOAFDKLAFSDGYAKYVTEATWEDDEKP
TCTKSTYKINWLGKGIITADKLVTSFNATETADAAGVELDVTFRAGKIEGIVN
GMDLEWNPDKTDFSLVSDNSVYVYAKAAKEALQAEGLPVDPTAPLFAFAGLIEE
CGVDIILALPKILATPKVQIAILGTGKAAYEKLVNAIGTKYKRAKGVYKFSAPLA
HMLTADADFMVLPSPFEPFGIQLHAMHYTPVAVSTGVLDTVKEGVTGFHMGALN
PDKLDEADADALAAVTRASEVFRAGRYPEMVAICISODLSWSIPAOKWEGILLIEVY
GKGVATAKKEIKVPAEKTGDLPAVSYAPNTLLKPVASVEGNGAARAPKVTITAPA
MGANRATTPSGSPSAATPKVTYTKPALPAIPAKPTAGLKLAGEASTTSTSENGAAS
GNGNGASASKISAAKPLVSAATKSA*
BASE COUNT 1073 a 1816 c 1852 g 1115 t
ORIGIN
Query Match 45.3%; Score 962.2; DB 8; Length 5856;
Best Local Similarity 86.2%; Pred. No. 5.2e-87;
Matches 1154; Conservative 0; Mismatches 8; Indels 177; Gaps 1;
QY 963 CACCGTCATCGCGCCCAAGGCGCAITGAGGCAITGTGAACGGCATGGACATGAGGAGTG 1022
Db 3570 CTCCTCCCGCAGCGCCCAAGGCGCAITGAGGCAITGTGAACGGCATGGACATGAGGAGTG 3629
QY 1023 GAACCCCAAGACCGACAGTTCTCTGTGCGCCCTACGACACAGACAGCGTCTACGCCGG 1082
Db 3630 GAACCCCAAGACCGACAGTTCTCTGTGCGCCCTACGACACAGACAGCGTCTACGCCGG 3689
QY 1083 CAAGCGCGCGCCCAAGGAGGCGCTCGAGCGGAGCTGGCGCTGCTGTGGACCCCAACCGC 1142
Db 3690 CAAGCGCGCGCCCAAGGAGGCGCTCGAGCGGAGCTGGCGCTGCTGTGGACCCCAACCGC 3749
QY 1143 CCCCGCTGTGCGCTTCATCGCGCGCGCTGGAGGAGCAGAAAGGCTGTGACATCATCTGCG 1202
Db 3750 CCCCGCTGTGCGCTTCATCGCGCGCGCTGGAGGAGCAGAAAGGCTGTGACATCATCTGCG 3809
QY 1203 CGCGCTGCCCAAGATCTCTGCGCCACCCCAAGTGCGATCGCCATCTCTGGGTACCGGCA 1262
Db 3810 CGCGCTGCCCAAGATCTCTGCGCCACCCCAAGTGCGATCGCCATCTCTGGGTACCGGCA 3869
QY 1263 GGCGCGCTTACCAGAGCTGTGTGAAGCGCATCGGCACCAAGTACAGGCGCGCGCGCAAGGG 1322
Db 3870 GGCGCGCTTACCAGAGCTGTGTGAAGCGCATCGGCACCAAGTACAGGCGCGCGCGCAAGGG 3929
QY 1323 CGTGGTCAAGTCTCGCGCGCGCGCTGGCGCACATGCTACCGCGCGCGCGCGCTCATGCT 1382
Db 3930 CGTGGTCAAGTCTCGCGCGCGCGCTGGCGCACATGCTACCGCGCGCGCGCGCTCATGCT 3989
QY 1383 GGTGCGCTCGCGCTTCGAGCGCGCTCGCGCGCTGATCAGTGTCACGCGCTACGCTACGCTAC 1442
Db 3990 GGTGCGCTCGCGCTTCGAGCGCGCTCGCGCGCTGATCAGTGTCACGCGCTACGCTACGCTAC 4049
QY 1443 CGTGGCGGTGTAGCTTCCACCGCGCGCTGTCGACACCGCTCAAGGAGGCGCGTACCGG 1502
Db 4050 CGTGGCGGTGTAGCTTCCACCGCGCGCTGTCGACACCGCTCAAGGAGGCGCGTACCGG 4109
QY 1503 CTTCCACA----- 1510
Db 4110 CTTCCACAGTGAGCGGGCTTCGGTGTGTGCTGTGAACITGTGCGCCCTGAAGTGTGCTGT 4169
QY 1511----- 1510
Db 4170 GCGGGGCGCTTTCGCGGGCTTCCTGACGCTGCTACCCCTGCTGCGCGCGCGCGGAAACC 4229
QY 1511----- 1510
Db 4230 ACCGGTATCGAAGACACCGCTGGCGCATGTGACGCGCTCCTCTTTTCCCGCTGTTCCCTGAT 4289
QY 1511-----TGGCGCGCTGAACCCCGACAGCTGACAGGCTGACGCGCGCGCGCGCGCGCGCG 1565
Db 4290 CGCAGTGGGCGCGCTGAACCCCGACAGCTGACAGGCTGACGCGCGCGCGCGCGCGCGCGCG 4349
QY 1566 CACCGTGGCGCGCTGACGCGCGCGCTGTTTGGCGCGCGCGCTACCCCGAGATGTTGGCCAA 1625
Db 4350 CACCGTGGCGCGCTGACGCGCGCGCTGTTTGGCGCGCGCGCTACCCCGAGATGTTGGCCAA 4409

```



```

QY 1626 CTGCATCAGCAGGACTGCTCTGTCCTCAAGCCGCCGACAACTGGGAGGCGCTGCTGGA 1685
Db 4410 CTGCATCAGCAGGACTGCTCTGTCCTCAAGCCGCCGACAACTGGGAGGCGCTGCTGGA 4469
QY 1686 GGAGTGGTGTACGGCAAGGCGCGTGGCCACCGCCCAAGAGGAGGAGATCAAGTGGC 1745
Db 4470 GGAGTGGTGTACGGCAAGGCGCGTGGCCACCGCCCAAGAGGAGGAGATCAAGTGGC 4529
QY 1746 CGTTGCGGAGAAATCCCGGCGACCTGCGCGCGTGTCTTACGCCCCCAACACCTTGAA 1805
Db 4530 CGTTGCGGAGAAATCCCGGCGACCTGCGCGCGTGTCTTACGCCCCCAACACCTTGAA 4589
QY 1806 GCCCGTGTCCGCTCCGTCGAGGCAACGCGCGCGCCGCAAGTCCGCGCACACCGC 1865
Db 4590 GCCCGTGTCCGCTCCGTCGAGGCAACGCGCGCGCCGCAAGTCCGCGCACACCGC 4649
QY 1866 CCCCAGGAGGCGCGTGGCGCGGACACCCCTCGGGCGCCCTGCGCGCGCGCGCCAC 1925
Db 4650 CCCCAGGAGGCGCGTGGCGCGGACACCCCTCGGGCGCCCTGCGCGCGCGCGCCAC 4709
QY 1926 CCCCAGGAGGCGCGTGGCGCGGACACCCCTCGGGCGCCCTGCGCGCGCGCGCCAC 1985
Db 4710 CCCCAGGAGGCGCGTGGCGCGGACACCCCTCGGGCGCCCTGCGCGCGCGCGCCAC 4769
QY 1986 CTTCAAGTGGCGGTCGAGGCTCCACACCTCGACCTCGAGAACCGCGCTGCTCCAA 2045
Db 4770 CTTCAAGTGGCGGTCGAGGCTCCACACCTCGACCTCGAGAACCGCGCTGCTCCAA 4829
QY 2046 CGGCAAGGCAAGGTCGCTGGCGCTCCAGACCTCGGCTGCGCAAGCCCTGCTCGGC 2105
Db 4830 CGGCAAGGCAAGGTCGCTGGCGCTCCAGACCTCGGCTGCGCAAGCCCTGCTCGGC 4889
QY 2106 CGCACCCGCAAGTCCGCC 2124
Db 4890 CGCACCCGCAAGTCCGCC 4908

RESULT 8
AF286320 2289 bp DNA linear PLN 30-OCT-2000
LOCUS Triticum aestivum granule bound starch synthase I (gbss1) gene,
DEFINITION complete cds.
ACCESSION AF286320
VERSION AF286320.1 GI:11037535
KEYWORDS Triticum aestivum.
SOURCE Triticum aestivum.
ORGANISM Triticum aestivum.
REFERENCE 1. 2289
AUTHORS McCue, K.F., Hurkman, W.J., Tanaka, C.K. and Anderson, O.D.
TITLE Starch Branching Enzymes Sbe1 and Sbe2 from Wheat (Triticum
aestivum cv. Cheyenne): Molecular Characterization, Developmental
Expression, and Homolog Assignment by Differential PCR
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2289)
AUTHORS McCue, K.F. and Anderson, O.D.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2000) United States Department of Agriculture,
Agricultural Research Service, 800 Buchanan Street, Albany, CA
94710-1105, USA
FEATURES
Source Location/Qualifiers
1. 2289
/organism="Triticum aestivum"
/cultivar="Cheyenne"
/db_xref="taxon:4565"
/tissue_type="endosperm"
204..2021
/gene="Gbss1"
204..2021
/gene="Gbss1"

```

```

/EC_number="2.4.1.21"
/note="glucosyltransferase; ADP glucose glucosyl
transferase; waxy; Wx"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AG27624.1"
/db_xref="GI:11037535"
/translation="MAALVTISLATSGTGLITDRFRAGFQVRRSPADAPLMRT
TGASAAPKQSRKRAHRTIRCLSMVWATSGAGMNLVFGAEMAPSKYKGGGLVIG
LPPMAANGHRVMIISPRYDQKADMTSVSEIKVADEYERVYKCYKRGVDFV
DHPCTFLKVRGKTKETIYGDADCTYEDNOLRSLKQAALAPRILDLNNPYSFGP
YGEDVYFVNCNDRHTGLIACYSNYSGLYRTAKVAFCHINISYGRSFSDFAOLN
LPDRFKSFDIDGDKPVEGRKLNWKGALQADKVLVSPYIAELISGEARGCEL
DNTRLITGITVINGMDVSEWDPAKFLAANDVITALEGKALNEALQAEVGLPVD
RKVLVAFITGRLEPEQKPDVMAAIEPILKEEDVQIVLLGTGKKKRELERLKSVEEKF
SKRVAVRENAPLAHOMAGADVLAITSFPCGLIQLOGMRYGTGCACATGGTGLDT
IMBGKTFEHMGLHSLVDCNVVVEPADVKVVTTLKRAVKVGTTPAYHEMVKNMCTQDLISW
KGPAAKWRDVLLELGVGSEPGVIGIEIAPLAMENVAAP"
BASE COUNT 480 a 684 c 735 g 390 t
ORIGIN
Query Match 25.3%; Score 537; DB 8; Length 2289;
Best Local Similarity 61.8%; Pred. No. 8.4e-45;
Matches 1011; Conservative 0; Mismatches 530; Indels 96; Gaps 6;
QY 171 CGCGCTGGACATCGTGTGCTGTGAGTGCCTTGGTCCCAAGACGGGGCGCT 230
Db 431 CGGCATGAACCTCGTTCGTGGCGCGAGATGGCGCCTGGAGCAAGACGGCGCT 490
QY 231 GGGCGATGTACTGTGCTGCTGCTATTTAGCTGTGCTCAAGCGGCCACCGCTCATCAC 290
Db 491 CGGCGAGCTCTCGGGGGCCTCCCGCCAGCATGCGCCGCAACGGTACCGGGTCTGGT 550
QY 291 CATTCGCCCTCGCTACGACCACTAGCTGACGCTGGACACCTCGGTGGTGGAGCAT 350
Db 551 CATCTCCCGCGCTACGACCACTAGAGGAGCGCTGGGACACCAAGCTGCTCGGAGAT 610
QY 351 CATGG-----CGGAGAGTCCGCTACTTCCACTCCATCAAGAGGGCGTCA 398
Db 611 CAAGGTCCGCGACGAGTACGAGAGGTGAGTACTTCCACTGCTCAAGCGCGGGGTGGA 670
QY 399 CGGCTGTGGATTGACACCCCTGCTTCTGCGCAAGTCTGGGCAAGACCGGCTCCAA 458
Db 671 CGCGTGTCTGTCGACACCCGCTTCTCTGGAGAGGTCCGGGCAAGACCAAGAGAA 730
QY 459 GCTGTACGGCCCCCGCTCCGCGCTGACTACTGTGACAAACCAAGCGCTTCGCCCTGTT 518
Db 731 GATCTACGGCGCGATGCGGCGAGCTAGGAGCAACCAAGCTTACGCTTACGCTGCT 790
QY 519 CTGCNAGCGCGCTATTGAGGCTGCGCGTGTGCTGCTC----- 558
Db 791 CTGCGAGCGCGCTTTGAGGCGCGCGGCTGCTGCTGCTCAACACACCCCTACTTCTC 850
QY 559 ---GGCGCGCGGAGGACTCGTCTGCTGGCCACGACTGGCGACTCCGCCCTCGGCTGCC 614
Db 851 CGGACCTTACGGGAGAGCGTGTGCTGCTGTCACGACTGGCACAGGCGCTTCTGGC 910
QY 615 CGTCTGTCTGAAGGACGAGTACCGCCCAAGGCGCAGTTTACCAAGGCGCAAGTCGGTCT 674
Db 911 CTGCTACCTCAAGAGCAACTACCACTGCTATATAGGACGCGCAAGGTAGCGTT 970
QY 675 GCATATCCACACATCGCTTCCAGGCGCGCATGTGGGAGGAGGCTTCAAGGACACGAA 734
Db 971 CTGCATCCACACATCTCGTATFAGGCGCGCTTCTCTTCTGACGACTTCGCGCAGTCA 1030
QY 735 GGTGCGCGCGCGCGCTTTTGCAAGCTGTGCGGACGCTATGCCAGGCTTTACAC 794
Db 1031 CCTGCCGACAGGTTCAAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
QY 795 TGAGGCCACCCCATGGAGGAGGACGAGAACGCCCGCTGACGCGGGAAGACCTACAGNA 854
Db 1083 -----CGGCTGGAGGGCGC-----AA 1099

```


QY	855	GATCAACTGGCTGAAGGTGGCAATATATCGCCCGACAAAGCTGTGACTGTGTGCCCCAA	914
Db	1100	GATCAACTGGATGAAGCGCGGATCTCGACGCGACAAAGTGTCTACGGTGGAGCCCTA	1159
QY	915	CTACGCGACCGATCGCTCGCGATGCGCGCGCGCGGTGTGAGCTGGACACCGTCAATCCG	974
Db	1160	CTACGCGGAGGAGCFCATCTCCGCGAAGCCAGCGGCTGCGAGCTCGACAAATCATGCG	1219
QY	975	CGCCAAAGCGATTGAGGCGATTGTGAACCGCATGACATTTGAGGAGTGAACCCCAAGAC	1034
Db	1220	CTCAGCGGATCACCGGATCTGCAAGCGCATGAGACGTGAGAGTGGGACCCGCGCAA	1279
QY	1035	CGACAAGTTCCTGTGCGCCCTACGACCAAGACAGCGTCTACCGCGCAAGCGCGCGC	1094
Db	1280	GGACAAGTTCCTGCGCGCAACTACGACCTCACACCGCTTGGAGGGAAGCGCTGAA	1339
QY	1095	CAAGGAGCCCTGAGCGGAGCTGGCGCTGCTGTGGACCCACCGCCCGCTGTGTGCG	1154
Db	1340	CARGAGGCGCTGACGCGGAGGTGGGCTGCCGCTGACCGGAAGTGGCCCTGGTGGC	1399
QY	1155	CTTCATCGCCCGCTGGAGGACGAGAGGTGTGGACATCATCTCGCGCGCCCTGCCCAA	1214
Db	1400	CTTCATCGCGAGCTGGAGGACGAGAGGCGCCGACGTGATGTCGCCGCTCCCGGA	1459
QY	1215	GATCCTGGCCACCCCGAAGGTGAGATCGCCATCTCTGGGTACCGGCAAGCGCGCTACGA	1274
Db	1460	GATCTTGAAGGAGGAGGACGCTCCAGATGTTCTCTCGGCAACCGGAAGAAGTTGA	1519
QY	1275	GAACTGTGTGAACGCAATCGCACCACAGTACAAGGCGCGCGCTTCATGCTGTGTCCTCGC	1334
Db	1520	CGCGTCTCAAGAGCTGGAGGAGAGTTCCTCCGAGCAAGGTGAGGCGCGTGTGTCAGGT	1579
QY	1335	CTCGGCGCCCTGCGGCACATGCTCACCGCGCGCGCGCTTCATGCTGTGTCCTCGC	1394
Db	1580	CAACGCGCGCTGCTCACAGATGATGCGCGCGCGCGAGTGTCTCGCGCTACCGACCG	1639
QY	1395	CTTGAGCCCTGCGCGCTGATCCAGCTGACGCGCAGCCATGACGCTACCGTGCCTGCTGT	1454
Db	1640	CTTCAGCCCTGCGCGCTCATCCAGCTCCAGGGGATGCGCTACCGGAACGCGCTGCGGTG	1699
QY	1455	AGCTTCCACCGCGCGCTGTGTCGACACCGTCAAGGAGGCGTCAACCGCTTCCACATGG	1514
Db	1700	CGCGTCCACCGCGCGCTGTCGACACGATCATGAGGCGCAAGACCGGTTTCCACATGG	1759
QY	1515	CGCCCTGA-----ACCCCGACAAGCTGGACGAGGCTGACGCGCAGCCCTTGGCGCCAC	1568
Db	1760	CCAGCTCAGCGTGCAGCTGCAACAGTGGTGGAGCGCGCGGACGTGAAGAAGTGTGGTACCAC	1819
QY	1569	GTGCGCGCTGCCAGCGAGTGTGCGGCGCGCGCTACCCCGAGATGTGTGGCCCACTG	1628
Db	1820	CCTGAAGCGCGCTCAAGGTGCTGCGGACGCCAGCTACCATGATGTTGTCAGAAGTGT	1879
QY	1629	CATCAGCGAGACCTGCTGCTGCTGCTCAAGCGCGCGCGCAAGGTGGAGGCGCTGCTGGAGGA	1688
Db	1880	CATGATCCAGATCTCTCTGGAAGGCGCGCGCAAGACTGGAGGAGCTGCTTCTGGA	1939
QY	1689	GTGCGTACGCGAAGGCGCGCTGGCCACCGCGCAAGAGGAGATCAAGTGTGCCGT	1748
Db	1940	ACTGGG---GTCAGGAGGAGCGAGCGAGGGGTCTATCGGCGAGGAGATGCGCGCTGCG	1996
QY	1749	TGCCGAGAAGATCCCG 1765	
Db	1997	CATGAGANCTGCGCG 2013	
RESULT 9	AF486514	1812 bp	linear
LOCUS	Hordeum vulgare cultivar oderbrucker granule bound starch synthase		
DEFINITION	I mRNA, complete cds.		
ACCESSION	AF486514		
VERSION	AF486514.1		
KEYWORDS	GI:21667427		

```
QY 559 -----GGCCCCGGGAGGACTGGCTTCTGCTGGCCAAACACTGGCACTCGGCCCTGGTGC 613
Db 641 CTGGTCCCTACGGGAGAGAGCTGGTTCGTGTGCAACGACTGGCACAGGGCCCTCTGG 700
QY 614 CGTCTGCTGAAGAGCAGAGTACCAAGCCCAAGGGCCAGTTCCACCAAGGCCAAGTCGGTGC 673
Db 701 CTTGTACCTCAAGAGCAACTACCAAGTCAATGGCATCTACAGGACGGCCAAAGGTGGCT 760
QY 674 TGGCTATCCACACATCGCTTCCAGGGCCGCGATGTGGGAGGAGGCTTTCACAGCACGA 733
Db 761 TCTGATCCACACATCTCTGACAGGGCCGCTTCTCTCGACGACTTTGGCGACGCTCA 820
QY 734 AGTGGCCCCAGGCGCTTTTGCAAGCTGGCTTCTCGGACGGCTATGCCAAGGTTTACA 793
Db 821 ACCTGCCCCGACAGGTTCAAGTCGCTCTCGACTTCAITGACGGCTACGACAAG----- 873
QY 794 CTGAGGCCACCCCATGGAGGAGGAGGAGAGAGGCCGCCCTGACGGGAAGACCTACAAGA 853
Db 874 -----CCGTTGAGGGGCGC-----A 889
QY 854 AGATCAACTGGCTGAAGGTTGGCATTATCGCGCCGACAAAGCTGGTGACTGTGTCGCCCA 913
Db 890 AGATCAACTGGATGAAGGCCGGGATCTCGAGGCCGACAAAGTGTCTGACGGTAGCCCT 949
QY 914 ACTACCGCACCGAGATCGCTGCGGATCGCGCGCGGCTGTGGAGCTGGACACCGTCATCC 973
Db 950 ACTACGCTGAGGAGCTCATCTCTGGCGAAGCCAGGGGCTGCGAGCTCGACAACTCATGC 1009
QY 974 GCGCCAAAGGCGATGAGGCGATGTGAACGGCATGTGACATGAGGAGTGAACCCCAAGA 1033
Db 1010 GCGTCAAGGGGATACCGGAATGTCACAGGCAATGACGCTGAGTGGGACGCCACCA 1069
QY 1034 CGACAAAGTTCCTGTGCGCCCTACGACCAAGCAAGCGCTCTACGCGCGCAAGGCCGCG 1093
Db 1070 AGGACAAGTTCCTGCGCGTCACTACGACATCACACCGCGTTGGAGGCCAAGCGCTGA 1129
QY 1094 CCAAGAGGCCCTTGCAGGCGAGCTGGGCTGCTGTGTGACCCCAAGCGGCCCTGTTCG 1153
Db 1130 ACAAGAGGCGCTGCAGGCGGAGTGGGCTGCGCGTGGACAGGAAGTGCCTGCTGGT 1189
QY 1154 CTTTCATCGCCGCTGGAGGAGCAAGAGGTTGACATCATCTCTGGGCGGCGCTGCCCA 1213
Db 1190 CTTTCATCGCAGCTGGAGGAGCAAGAGGCGCCGACGCTGATGATCGCGGCTATCCCG 1249
QY 1214 AGATCTGCGCACCCCAAGGTGCAGATCGCCATCTCTGGGTACCGGCAAGGCCGCTACG 1273
Db 1250 AGATCTGAGGAGGAGGCTCCAGATCATCTCTCTGGCACCGGGAAGAAGTTTG 1309
QY 1274 AGAAGCTGTTGAAGCGCATGGCACCAAGTACAAAGGCGCGCCCAAGGCGGTGGTCAAGT 1333
Db 1310 AGAAGCTGCTCAAGAGCATGGAGGAAGTTCCCGGCAAGGTGAGGCGCGTGGTCAAGT 1369
QY 1334 TCTCGGCGCCCTGGCGCATCTCACCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1393
Db 1370 TCAACGCGCGCTAGCTACACAGATGATGCGCGCGCGCGGCTGCTGCTGCTGCTGCTG 1429
QY 1394 GCTTCGAGCCCTGCGGCTGATCCAGCTGACGCCATGACGCTACGCTACGCTGCGCGCTG 1453
Db 1430 GCTTCGAGCCCTGCGGCTCATCCAGCTCCAGGGAATGCGCTATGNAACGCGCTGGTGT 1489
QY 1454 TAGCTTCCACCGCGGCTGCTGTCGACACCGCTCAAGGAGGCGCTCACCGGCTTCCACATGG 1513
Db 1490 GCGGCTCCACCGCGGCTGCTGACACGATGCTGGAGGCAAGACCGGCTTCCACATGG 1549
QY 1514 GCGCCCTGA-----ACCCGCAAGCTGACGAGGCTGACCGCGGCTGCGGCGCCCA 1567
Db 1550 GCGGCTTACGCTGCTCAAGCTGTTGGAGCGCGGAGCTGGAAGAGTGGCGGACCA 1609
QY 1568 CGGTGCGCGCTGCGGAGGAGTGTTCGCGGCGCGGCTGCTACCGCGGAGTGGTGGCGCAACT 1627
Db 1610 CCGTGAAGCGGCGCTCAAGGTCGTGCGGCACGCGCGGCTGACGAGAGTGGTCAAGAACT 1669
```

```
QY 1628 GCATAGCCAGGACCTGTCTCTGTTCCAAAGCCGCCAGAAAGTGGAGGCGCTCTGAGG 1687
Db 1670 GCATATCCAGGATCTCTCTGGAAGGAGCTGCCAAGAACTGGGAGGACGTGCTTCTGG 1729
QY 1688 AGTGTGTACGCAAGGCGGCTGGCCACCCGCCAAGAGGAG 1731
Db 1730 AACTGGGGTGGAGGAGCGAGCGGGGATCGTCGCGGAGGAG 1773

RESULT 10
HWAXYR
LOCUS HWAXYR HWAXYR 2311 bp mRNA linear PLN 12-SEP-1993
DEFINITION Barley mRNA pcwx27 for waxy locus.
ACCESSION X07932
VERSION X07932.1 GI:19128
KEYWORDS starch synthase; waxy locus.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 2311)
AUTHORS Zohde W
TITLE Direct Submission
JOURNAL Submitted (21-JUN-1988) Zohde W., Max Planck Institute, Erwin Baur
Institut, D-5000 Koln 30, FRG
REFERENCE 2 (bases 1 to 2311)
AUTHORS Rohde W., Becker, D. and Salamini, F.
TITLE Structural analysis of the waxy locus from Hordeum vulgare
JOURNAL Nucleic Acids Res. 16 (14B), 7185-7186 (1988)
MEDLINE 88303345
PUBMED 2970062
COMMENT The sequence overlaps with the genomic sequence of the waxy locus,
x07931, from bp 1918 to bp 2301.

FEATURES
source
1..2311
/organism="Hordeum vulgare subsp. vulgare"
/strain="vogelsanger Gold"
/db_xref="taxon:112509"
/map="chromosome 1"
/clone="pcwx27"
/clone_lib="EMBL4"
226..2037
/note="starch synthase"
/codon_start=1
/protein_id="CAA30756.1"
/db_xref="GI:19129"
/db_xref="SWISS-PROT:P09842"
/translation="MAALATSLATSGTLVGTDRFRFGFQGLRPRNPADALGNRT
IGASAAPKQSRKAHGRSRLSVVVSATSGMNLVFGAEMAPWSKTIKGLDVLGLP
PAMAANGHRVMVSPRYDQYKDAWDTSVSEIKVADEYERVFFHCYKRGVDFIDH
PWFLKVRGKTKEIKYDPADIDYEDNOQRFSLCQAALAPRILNINNNPYSGPYG
EDVFCVNDWHTGLLACYLKSNTSGNIGYRTAKVAFCHINISYQGRFSPDFQAQLNP
DRKSSDFDIDGDKPEGRKLNMMKAGILOADKVLTVSPYAFELISGEARGCLDN
IMRLGITGIVNGMVSNDPKDFLAVNYDITALEAKALKEALQEVGLPVDNR
VPLVAFNAPLAPMMAGADLLAVTSRFFCGILQGMRYGTPCVCASTGGGLDFTIV
EGKTFHMGRLSDVCNVPEPADVKVATTKIKRAVKVYGTPAYQEMVKMCMQIDLSWK
PAKNWEDVLLELGVGSEPGIVEIAPLAVMENVAAP"
2276
/polyA_site
/note="poly A site (pcwx4)"
2283
/polyA_site
/note="poly A site (pcwx27)"
2301
/polyA_site
/note="poly A site (pcwx18)"
BASE COUNT 501 a 688 c 714 g 403 t
ORIGIN

Query Match 25.0%; Score 531.2; DB 8; Length 2311;
Best Local Similarity 61.6%; Pred. No. 3.1e-44;
Matches 988; Conservative 0; Mismatches 523; Indels 93; Gaps 5;

QY 170 GCGCGTGGACATCGTATGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGGCGCC 229
```


AUTHORS Patron,N.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK
FEATURES Location/Qualifiers
source 1. .1827
/organism="Hordeum vulgare"
/cultivar="SB 85750"
/db_xref="taxon:4513"
1. .1827
/product="granule bound starch synthase I"
/protein_id="AA074051.1"
/db_xref="GI:21667436"
/translation="MAALATSQLATSGVLGVTRDSAPSMFRHAGFQGLRPRKPADGT
FGMRTVGAAPQSRKAHRNRCISVVRATSGMNLVFWGAEAPWSKTLGLGVD
LGGLPMAARQSRVWVSPRYDQYKADMTSVISEIKVADEYERVFHCYKRGVDR
VFTDHPWLEKVRGKTEKIYDPDAGTDVEDNQRFSLICQAALEAPRILNLNNPYE
SGPYGDEVFVNCNMTGLLACVLKSNYQSNGLYRTAKVAFCHLNIYSQGRSFDFPA
QLNLPDRFSSPFDIGDKPVEGRKLNKMKALQAKVLTSPSYAEELISGEAFG
CEDLNRLGFIITGVNMDVSEWDPDKFLAVNYDITITALEKALNKELALQAEVL
PVDKPLVAFIRLEQKGPDMVIAAIPELKEEDVQIILLGTGRKKFKLKSHEE
KFPKRVYRNFAPLAHOMNAGADLLAVTSREPGLIOLQGMRYGTPCVCASTGSL
VDTPIVEGKTFHMRSLSDVNCNVVEPADVKVATLKRKVVGTAYQEMVKNQMID
LSWKGPAAKNWEDVLLLELGVGSPGIVGEIAPLAVENVAAP"
BASE COUNT 364 a 567 c 602 g 294 t
ORIGIN

Query Match 24.9%; Score 529.6; DB 8; Length 1827;
Best Local Similarity 61.5%; Pred. No. 4.8e-44;
Matches 987; Conservative 0; Mismatches 524; Indels 93; Gaps 5;
QY 170 GCGGCTGGACATCGTATGTTGCTGCTGAGTGGCCCTTGTCCTCAAGACGGGCGGCC 229
DB 236 GCGGCTGAACCTCGTGTCTGCGGCCCGAGATGGCGCCCTGGAGCAAGACCGCGGCC 295
QY 230 TGGCGGATGTGACTGTGGGCTGCTGCTATTGAGCTGTGTCAGGCGGCCACCGCGTCTATGA 289
DB 296 TTGGCGAGTCTCGGGGCTTCCACCGCATGGCGGCCAAGCGGTACCGGGGTATGG 355
QY 290 CCATGTCCTGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGTGCTGGACA 349
DB 356 TCGTCTCCCGCGCTACGATCAGTACAGGACGCTGGGACACCGCTCATCTCCGAGA 415
QY 350 TCATGG-----GGAGAGTCCGCTACTTCCACTCCATCAAGAGCGGTGC 397
DB 416 TCAAGTTCGTGACGATGACGAGGTGAGGTCTTCCACTGCTACAAGCGGAGTGG 475
QY 398 ACCGCGTGTGATGACACCCCTGTTCTCTGGCCAAAGTCTGGGCAAGACCGGCTCCA 457
DB 476 ACCGCGTGTGATGACACCCCTGTTCTCTGGAGAGGTCCGGGCAAGACCAAGAGA 535
QY 458 AGTGTACGCGCCCGCTCGCGCGCTGACTACCTGGACAAACCAAGCGCTTCGCCCTGT 517
DB 536 AGATCTACGCGCCCGAGCGCGGACGACTACGAGGACCAACCAAGCGCTTCAGCCTTC 595
QY 518 TCTGCAAGCGCGCTATTGAGCTGCCCGCGTGTGCGCTTC----- 558
DB 596 TCTGCCAGGCGCGCTCGAGGCAACCCAGGATCTCTCAACCTCAACCAACCCCTACTTTT 655
QY 559 -----GSCCGCGGAGGATGCTGCTCTTCTGGCCAAAGCTGSCACTCCGCGCTGTCG 613
DB 656 CTGTCCTTACGGGAGACGTTGTGTGTGTGCAACAGTGTGCAACGCGGCGCTTCTGG 715
QY 614 CCCTCTCTGCTGAAGAGGAGTACAGCCCAAGGCGCAGTTTACCAAGGCGCAAGTGGTGC 673
DB 716 CCGTCTACCTCAAGAGCACTACCAAGTCCATGGCATCTACAGGACGCGCAAGGTGCGCT 775
QY 674 TGGCTATCCACATCGCTTCCAGGCGCGCATGTGGGAGGAGGCTTCAAGSACACGA 733
DB 776 TCTGCATCCCAACATCTCTACGAGGCGGCTTCTCTTCTCGACGACTTTTCGCGAGCTCA 835

QY 734 AGCTGCCCGGAGCGCGCTTTGACAAAGCTGGCTTCTTCGACGGCTATGCCAAGGTTTACA 793
DB 836 ACCTGCCCGGACAGGTTCAAGTCGCTCTTCGACTTCATTGAGGGCTACGACAAG----- 888
QY 794 CTGAGGCCACCCCATGAGGAGGAGGAGAGAGCCCGCTGACGGGAAAGACCTPACAAGA 853
DB 889 -----CCGTTGAGGGGCGC-----A 904
QY 854 AGATCAACTGGCTGAAGGTTGGCATTTATCGCGCGGACAAAGCTGGTACTGTGCGCCA 913
DB 905 AGATCAACTGGATGAAGGCGGGATCTCGAGCGGACAAAGTGTGACGCTGAGGCCCT 964
QY 914 ACTAGCGACCGAGATCGCTCCGATGCGCGCGGCTGGAGTGTGACACCGCTCAATCC 973
DB 965 ACTAGCTGAGAGCTCATCTCTGCGAAGCAGCGGCTGCGAGCTGACACATCATGTC 1024
QY 974 GCGCAAGGGCATTTGAGGGCATTTGTAACGGCATGGACATGAGGAGTGAACCCCAAGA 1033
DB 1025 GCCTCAGGGGATCACCGGAATCTCAACGGCATGGAGCTCAGTGTGAGGAGGAGGAGG 1084
QY 1034 CCGAAAGTTCTCTGTCGCGCTTACGACCAAGAGCTGTACGCGGCAAGGCGCGCG 1093
DB 1085 AGGAAAGTTCTCTGCGCTCAACTACGACATCACACCGCTTGGAGGCCAAGGCGCTGA 1144
QY 1094 CCAAGGAGGCGCTGACGCGCGAGCTGGCCTGTGCTGTGGACCCACCGCCCTGTTTCG 1153
DB 1145 ACAAGGAGGCTGACGCGCGAGGTGGGCTGCGGCTGAGCAAGAGGTGCGCTGTTGG 1204
QY 1154 CTTTCATCGGCGCTGGAGGAGGAGGTGTGGACATCATCTGCGCGCGCTGCCCA 1213
DB 1205 CTTTCATCGGAGCTGGAGGAGGAGGAGGCGCGCGCTGATGATCGCGCGCTGCCG 1264
QY 1214 AGATCTTGGCCACCCCAAGGTGCAGATGCGCATCTCTGGGTACGCGCAAGGCGCGCTACG 1273
DB 1265 AGATCTTGAAGGAGGAGGAGCTCCAGATCATCTCTTGGCACCGGGAAGAAGTTTG 1324
QY 1274 AGAGCTGGTGAAGCCATCGGACCAAGTACAGAGGCGCGCGCAAGGCGCTGTCAAGT 1333
DB 1325 AGAAGCTGCTCAAGAGCATGGAGGAGAGTTCCCGGCGCAAGGTGAGGCGCTGTGTCAGGT 1384
QY 1334 TCTCGGCGCGCTTGGCGCACATGCTCACGCGCGCGCTTCTGCTGTGCTGCTCGC 1393
DB 1385 TCAAGCGCGCTAGCTACCCAGATGATGCGCGCGCGCTTGTGCTGCTGCTACCGAGCC 1444
QY 1394 GCTTCGAGCGCTGCGGCTGATCAGCTGACGCGCATGCTACGTTACCGTGGCGGCTGG 1453
DB 1445 GCTTCGAGCGCTGCGGCTCATCAGCTCCAGGAAATGCGTATGGAACGCGCTGCTGT 1504
QY 1454 TAGCTTCCACGCGCGCTGCTGCGACACCGTCAAGGAGGCGCTGACCGGCTTCCACATGG 1513
DB 1505 GCGGCTCCACGCGCGGCTGCTGCGACGATGCTGGAGGCGCAAGCTGCTTCCACATGG 1564
QY 1514 GCGCGCTGA-----ACCGCGACAAAGCTGGACGAGGCTGAGCGCGACGCGCTGCCGCA 1567
DB 1565 GCGCGCTCAGCGCTGCTGCAACGTTGTTGAGCGCGCGAGCTGAAGAGGTGGCGGACCA 1624
QY 1568 CCGTGGCGCGTGGCAGGAGTGTGTTGGGCGCGCGCTACCCCGAGATGTTGCGCAACT 1627
DB 1625 CCCTGAAGCGGCGCTCAAGTCTGTCGCGACGCGCGGCTTACCGAGAGTGTGCAAGACT 1684
QY 1628 GCATCAGCAGGAGCTCTCTGTTCCAAAGCGCGCGGAGGAGTGGAGGCGCTTGTGGAGG 1687
DB 1685 GCATGATCCAGGATCTCTCTGGAAGGAGCTTGCACAAAGTGGAGGAGCTGCTCTGG 1744
QY 1688 AGTGTGTACGCAAGGCGGCTGGCGCACCGCGCAAGAGGAG 1731
DB 1745 AACTGGGGTGGAGGAGGAGCGCGGATGCTGCGCGAGGAG 1788

RESULT 13
AF163319
LOCUS
DEFINITION

AF163319 2127 bp mRNA linear PLN 14-FEB-2000
Triticum aestivum granule-bound starch synthase I (wx-DI) mRNA,
DEFINITION

Db	1190	CTTTCATCGCAGCTGGAGGACAAAGGCGCCGACGTGATGTCGCGCCATCCGG	1249
QY	1214	AGATCTGGCCACCCCAAGTGCAGATCCCATCTCTGGTACCGGCAAGCCGCTACG	1273
Db	1250	AGATCTTGAAGGAGGAGGACCTCCAGTATCTCTTGGCACCAGGGAAGAAGTTG	1309
QY	1274	AGAAGCTGTGTGAAGGCATCGGCACCAAGTACAAAGGCGCGCCGCAAGGGCGTGTCAAGT	1333
Db	1310	AGAGCTGCTCAAGAGCATGAGGAGAGTTCGCGGCAAGGTGAGGGCGTGTGAGGT	1369
QY	1334	TCGCGGCGCCCTGGGCGCATGCTCACCGCGCGGCGCCGACTTCATGCTGGTGCCTTCGC	1393
Db	1370	TCAACGCGCGCTAGCTACAGATGATGGCGCGCGCGGACTTGCCTGCTCACCAGCC	1429
QY	1394	GCTTCGAGCCCTCGGCGCTGATCCAGCTGCACGCCATGCACTACGTTACCGTGCCTGG	1453
Db	1430	GCCTGAGCCCTCGGCGCTCATCCAGCTCCAGGAATGCGCTATGGAACCGCTGCGTGT	1489
QY	1454	TAGCCTCCACCGCGCGCTGGTGCACACCGTCAAGGAGGCGCTCACCGGCTTCCACATGG	1513
Db	1490	GCAGCTCCACCGCGCGCTCGTCGACAGATCGTGGAGGCAAGACCGGCTTCCACATGG	1549
QY	1514	GCSCCTGA-----ACCCGACAGCTGACGAGGCTGACGCGAGCCCTGCGCCGCA	1567
Db	1550	GCCGCTCAGCTGCACTGCAACCTGGTGGAGCGCGCGGACGTGAAGAAGGTGGCGACCA	1609
QY	1568	CCGTGCGCGCTGCGAGGAGTGTTCGCGGCGCGCGCTACCCGAGATGTGGCCAACT	1627
Db	1610	CCCTGAAGCGGCGCTCAAGTCTGCGGACGCGCGGCTACAGGAGATGTCAAGACT	1669
QY	1628	GCATGACCGAGGACCTGTCTGTGTCACAGCCCGCCAGAGTGGGAGGCGCTGCTGGAGG	1687
Db	1670	GCCTGTGCTCAGGATCTCTCTGGAAGGAGCTGCAAGAACTGGGAGGACGTCTTCTGG	1729
QY	1688	AGTGCTGTACGCAAGGCGCGCTGCGCCACGCGCCAGAGAGGAG	1731
Db	1730	ACTGGGGTGGAGGAGGAGCGCGGGATCGTTCGCGAGGAG	1773
RESULT 15			
AF486516			
LOCUS			
DEFINITION			
Hordeum vulgare cultivar Yon M kei non-functional granule bound			
starch synthase I, complete sequence.			
ACCESSION			
AF486516			
VERSION			
AF486516.1 GI:21667431			
KEYWORDS			
SOURCE			
Hordeum vulgare.			
ORGANISM			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Poideae; Triticeae; Hordeum.			
REFERENCE			
1 (bases 1 to 1827)			
AUTHORS			
Patron, N., Smith, A., Fahy, B., Hylton, C., Naldrett, M., Rossmagel, B.			
and Denyer, K.			
A mutation in the 5' non-coding region of the barley GBSSI gene			
alters its temporal and spatial expression and reduces GBSSI			
activity and amylose content in the endosperm			
Unpublished			
2 (bases 1 to 1827)			
JOURNAL			
Patron, N.			
REFERENCE			
Direct Submission			
AUTHORS			
TITLE			
Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,			
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK			
JOURNAL			
Location/Qualifiers			
FEATURES			
1..1827			
/organism="Hordeum vulgare"			
/db_xref="taxon:4513"			
misc_feature			
1..1827			
/note="contains non-functional granule bound starch			
synthase I due to premature stop codon"			

BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
985; Conservative			
24.8%; Score 526.4; DB 8; Length 1827;			
61.4%; Pred. No. 1e-43;			
0; Mismatches 526; Indels 93; Gaps 5;			
QY	170	GCAGCTGTGATCGTCAATGCTGCTGAGTGC3CCCTTGTGTCACAGCGCGCGCC	229
Db	236	CGGCAATGACCTGCTGCTGCGCGCGGAGATG3CGCCCTGGAGCAAGACCGCGGCC	295
QY	230	TGGCGCATGTGACTGTGGCTGCTATGATGATG3TCAAGCGCGCGCCACCGGCTCATGA	289
Db	296	TTGGCGAGCTGCTCGCGCGGCTTCCACCAAGCATG3CGCGCAAGCGTCAAGCGGTGATGG	355
QY	290	CCATGCCCCCTCGGTACGACGACGATGCTGACGCTCGGAGACCTCGGTGCTGACACA	349
Db	356	TCGTCTCCCCGCGCTACGATCAGTACAAAGGACGCTCGGACACCAAGGCTCATCTCCGAGA	415
QY	350	TCATGG-----GCGAGAAAGTCCGCTACTTCCATCCATCAAGAAGGCGCTGC	397
Db	416	TCAGGTGCTGACGAGTACGAGAGGTGAGGTTCTTCCACTGCTACAAGCGCGAGTGG	475
QY	398	ACGCGTGTGATGACACACCCCTGCTTCTGGGCCAAGGCTCTGGGGCAAGACCGGCTCCA	457
Db	476	ACGCGGTGTTATCGACACCCCGTGTCTCTGGAGAAAGTCCGGGGCAAGCAAGGAGA	535
QY	458	AGCTGTAGCGCCCGCTCGCGGCTGCTACCTGCTG3ACAAACCAAGCGCTTCGCCCTGT	517
Db	536	AGATCTAGGCGCCGACGCGGACGAGTACGAG3ACAAACCAAGTACGCTTACGCTTC	595
QY	518	TCTCAAGGCGCTATTGAGGCTGCCGCGCTGCTGCCCTTC-----	558
Db	596	CTGCCAGGAGCGCTCGAGGACACCCAGATCTCAACCTCAACAACAACCCCTACTTTT	655
QY	559	-----GGCCCCGGGAGGAGCTGCTCTCTGGGCCAAGGCTGCACTCCGCTTGGTGC	613
Db	656	CTGTCTCCTACGGGGAAGAGCTGGTGTCTGCTGTCACAGCTGGCAGCGGCTTCTGG	715
QY	614	CCGTCTCTGCTGAAGGAGGAGTACCAGCGCCAAAGGCGCTAGTTCAACCAAGCGCCAAAGTCGGTGC	673
Db	716	CTGCTACCTCAAGAGCAACTACCACTCCATGSCATCTACAGGACGCGCAAGGTGSCCT	775
QY	674	TGGTATCCAAACATCGCTTCCAGGCGCGCATG3GGGAGGAGGCTTTCAAGACACAGA	733
Db	776	TCTGATCCACACATCTCGTACCAAGGCGCGCTTCTCCTTCGACGACTTTTTCGACAGTCA	835
QY	734	AGCTGCCCGCGCGCTTTGACAAAGCTGGCTTCICGGAGCGGCTATGCCAAGGTTTACA	793
Db	836	ACCTGCCCGAGAGGTTCAAGTCTGCTCTCACTTCACTTGGCGCTACGACAG-----	888
QY	794	CTGAGGCGACCCCATGGAGGAGGACGAGAAAGCCCGCTGACGCGGAAGACCTTACAAGA	853
Db	889	-----CCGCTGGAGGGGCGC-----A	904
QY	854	AGATCACTGGCTGAAGGTTGGCAATTATCGCCGCGCCGACAGCTGGTACTGTGTCGCCCA	913
Db	905	AGATCACTGGATGAAGCGCGGATCTCTGACAGGCGCCACAAGGTTCTACGCTGAGCCCT	964
QY	914	ACTACGCGACGAGATCGTCCGATCGCGCGCGGCTGTTGGAGCTGGACACCGCTCATCC	973
Db	965	ACTAGCTGAGGAGCTCATCTCTGGCGAAGCCAGG3GCTGGAGCTGCACACATCATGC	1024
QY	974	GCGCCAAGGCGATTGAGGCGCATTTGAAACGSCATG3ACATTTAGGAGTGGAAACCCCAAGA	1033
Db	1025	GCCTCACGGGATCACCGGAATCGTCAACGSCATG3GAGTCAAGTGGAGGACCCACCA	1084
QY	1034	CCGACAAGTCTCTGCTGCGCCCTACGACACAGACGCGTCTAGCCGCGCAAGGCGCGCG	1093
Db	1085	AGGACAAGTCTCTCGCGCTCAACTACGACATCAACACCGGCTTGGAGGCCAAGCGGTGA	1144
QY	1094	CCAAGGAGGCGCTCAGGCGGAGCTGGGCGCTGTGCTGTGGACCCGACCCGCGCGCTGTGG	1153

Db 1145 ACAAGGAGGACATGACGAGCGAGGTGGGCTGCCGTGGACTGGAAGGTGCCGCTGGTGG 1204
QY 1154 CCTTCATCGGCGCCTGGAGGAGCAGAAGGGTGTGCACATCATCTGGCGCGCCTGCCCA 1213
Db 1205 CCTTCATCGGCGAGGTGGAGGAGCAGAAGGGGCCGCCACCTGATGTCGCCGTATCCCGG 1264
QY 1214 AGATCCTGGCCACCCGCCAGGTGCAGATGCCATCTCTGGTACCGCGCAAGGCGCCCTACG 1273
Db 1265 AGATCCTGAAGGAGGAGGACGTCACAGATCATCTCTGGCACCCGGGAAGAAGATTG 1324
QY 1274 AGAAGTGTGTGACGCCATCGGCACCAAGTACAAAGGCGCGCCCAAGGGCGTGGTCAAGT 1333
Db 1325 AGAAGTGTCTAAGACGATGAGGAGAAGTTCCTGGGCAAGGTGAGGCGCGTGGTCAAGT 1384
QY 1334 TCTGCGGCGCCTGGCGGCATGCTCACCGCGGCGCGGCGGCTCATGCTGTGGTCCCTCGC 1393
Db 1385 TCAACGCGCGGTAGTCTACCAAGATGATGGCGGCGCGGCTGCTCGCTGTCAACGAGCC 1444
QY 1394 GCTTCGAGCCCTGCGGCTGTATCCAGCTGCAGCCCATGCTACGTACGTACGTCGCCCGTGG 1453
Db 1445 GCTTCGAGCCCTGCGGCTGTATCCAGCTGCAGGAAATGGCTATGGAACGCCGTGGTGT 1504
QY 1454 TAGCCTCCACCGCGGCTGGTGTGACACCGTCAAGGAGGCGGTACCGGCTTCCACATGG 1513
Db 1505 GCGGCTCCACCGCGGCTGTGACACCGTGTGAGGAGGACCTGGTTCACATGG 1564
QY 1514 GCGGCTGA-----ACCCGACAGCTGGAGGCTGACGCCGACGCCCTGGCGGCA 1567
Db 1565 GCGGCTGACGCTGACTGCAACGTTGGTGGAGCGCGGACCTGAAAGAGGTGGCGACCA 1624
QY 1568 CGGTGCGCGTGCCAGCGAGGTGTTTGGCGGCGCGCTACCCCGAGATGGTGGCCAACT 1627
Db 1625 CCTGAGCGGCGGTCAAGGTGCTGGGACGCGCGCGTACCAGGAGATGGTCAAGAACT 1684
QY 1628 GCATCAGCCAGGACCTGTCTCTCCAGCCCGCCAGAGTGGGAGGCGCTGCTGGAGG 1687
Db 1685 GCATGATCCAGGATCTCTCTGGAAGGACCTGCCAAGAACTGGGAGGAGGTGCTTCTGG 1744
QY 1688 AGTGTGTACGGCAAGGCGGCTGGCCACCGCCCAAGAGGAG 1731
Db 1745 AACTGGGGGTGGAGGGAGCGAGCGCGGGGATGTCGCGCGAGGAG 1788

This Page Blank (uspto)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 05:08:24 ; Search time 308.686 Seconds
(without alignments)
15495.523 Million cell updates/sec

Title: US-09-980-771a-2

Perfect score: 2124

Sequence: 1 atggctgttgcctctaccag.....ccgcaccgcgaagtccgcc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	704	33.1	1696	22 AAC86955	CDNA encoding a gr
2	478	22.5	2267	18 AAX63355	Granule bound star
3	473.2	22.3	1915	19 AAX60319	DNA sequence of th
4	447.6	21.1	2542	19 AAV29753	Oryza sativa waxy
5	316.8	14.9	2161	24 ABK53210	Potato granule-bou
6	298	12.0	2115	21 AAC32824	Arabidopsis thalia
7	258	12.1	2183	24 ABK53215	Canna edulis granu
8	221	10.4	2807	21 AAC86435	Wheat starch synth
9	221	10.4	2842	21 AAC86411	Wheat starch synth

10	219.2	10.3	2097	19 AAV29755	Zea mays soluble s
11	218.2	10.3	2826	19 AAV01528	Wheat granule-boun
12	217.2	10.2	2423	19 AAV70958	DNA encoding maize
13	214	10.1	2107	21 AAC86412	Wheat starch synth
14	214	10.1	2107	21 AAC86434	Wheat starch synth
15	213.4	10.0	2380	19 AAV66834	Zea mays soluble s
16	212.4	10.0	2478	18 AAT67287	Soluble starch syn
17	212.4	10.0	9024	21 AAC86431	Wheat SSII gene SE
18	210.4	9.9	2920	24 ABK88112	CDNA encoding modi
19	209.2	9.8	2939	21 AAC86410	Wheat starch synth
20	209.2	9.8	2939	21 AAC86433	Wheat starch synth
21	208.8	9.8	2946	24 ABK88115	CDNA encoding barl
22	208.8	9.8	2951	24 ABK88114	CDNA encoding barl
23	208.2	9.8	2950	24 ABK88113	CDNA encoding barl
24	207	9.7	2248	21 AAZ50651	Corn starch synth
25	206.6	9.7	1798	21 AAZ50647	Corn starch synth
26	206.6	9.7	2019	21 AAZ50646	Corn starch synth
27	203.8	9.6	1724	24 ABA01821	Rice starch synth
28	161	7.6	2007	19 AAV70959	DNA encoding maize
29	161	7.6	2007	19 AAV29754	Zea mays soluble s
30	161	7.6	2085	18 AAT67286	Soluble starch syn
31	161	7.6	2085	19 AAV66833	Zea mays soluble s
32	151	7.1	1479	24 ABQ90258	M. capsulatus gene
33	151	7.1	1479	24 ABQ90406	M. capsulatus gene
34	125.6	5.9	333	21 ACN74569	Human ORFX ORF124
35	125.6	5.9	333	24 ABN18316	Human ORFX polynuc
36	118.2	5.6	2239	19 AAV01527	Wheat soluble star
37	118.2	5.6	2805	21 AAZ24487	Wheat soluble star
38	116	5.5	2662	20 AAX34651	CDNA sequence of w
39	110.8	5.2	65140	22 AAD17184	Streptomyces nous
40	110.8	5.2	125401	22 AAD17186	Streptomyces nous
41	110.6	5.2	2533	15 AQA45183	Soluble rice starc
42	108	5.1	1758	17 AAT32325	Soluble starch syn
43	106.8	5.0	2277	19 AAV13836	Humo sapiens mamma
44	106.8	5.0	2277	19 AAV05372	Human telomerase p
45	106.4	5.0	1984	21 AAC46797	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAC86955

ID AAC86955 standard; cdna; 1696 BP.

AC AAC86955;

XX 02-APR-2001 (first entry)

DT CDNA encoding a granule bound starch synthetase II (GBSSII).

DE Granule bound starch synthetase II; GBSSII; starch grain;

KW adenosine diphosphate glucose-alpha,4-glucan alpha4-glucosyltransferase;

KW starch synthetase; ds.

XX Chlamydomonas reinhardtii.

XX Key Location/Qualifiers

FT CDS 3..719

FT /*tag= a

FT /product= "granule bound starch synthetase II (GBSSII)"

XX FR2793806-AL.

XX 24-NOV-2000.

XX 21-MAY-1999; 99FR-0006494.

XX 21-MAY-1999; 99FR-0006494.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Dhulst C, Ball S;

XX WPI; 2001-052291/07.
DR P-PSDB; AAB31175.
XX
PT New recombinant nucleic acid encoding fusion of starch synthase and
PT second component, useful in pharmaceutical and food compositions, is
PT targeted to starch granules -
XX
XX
PS Claim 3; Fig 1; 52pp; French.
XX
XX The present sequence encodes a granule bound starch synthetase II
CC (GBSSI). The specification describes a recombinant nucleic acid,
CC comprising a polynucleotide encoding an adenosine diphosphate
CC glucose-alpha1,4-glucan alphas4-glucosyltransferase or starch
CC synthetase, placed upstream of a sequence that encodes a polypeptide
CC of interest. The adenosine diphosphate glucose-alpha1,4-glucan
CC alphas4-glucosyltransferase protein can migrate to sites of
CC biosynthesis of starch grains in plant cells, becoming associated
CC with these grains. The recombinant nucleic acid sequence is used to
CC target polypeptides of interest to starch grains.
XX
SQ Sequence 1696 BP; 347 A; 509 C; 571 G; 269 T; 0 other;

Query Match 33.1%; Score 704; DB 22; Length 1696;
Best Local Similarity 100.0%; Pred. No. 8e-94;
Matches 704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1421 TGCACGCCATGCACTACGGTACCGTACCGTGTAGCTCCACCGCGGCTGTGTGACA 1480
DB 13 TGCACGCCATGCACTACGGTACCGTGTAGCTCCACCGCGGCTGTGTGACA 72

QY 1481 CCCTCAAGGAGGCGGTACCGGCTTCCATGCGGCGCTTGAACCCGACAACTGAGC 1540
DB 73 CCCTCAAGGAGGCGGTACCGGCTTCCATGCGGCGCTTGAACCCGACAACTGAGC 132

QY 1541 AGCTGACGCCGCGCTGTGCGCGCCACCGTGTGCGCGTGTGCGGAGTGTTCGCGCG 1600
DB 133 AGCTGACGCCGCGCTGTGCGCGCCACCGTGTGCGCGTGTGCGGAGTGTTCGCGCG 192

QY 1601 GCGCTTACCCCGAGTGTGGCCCACTGCAATGACCCAGGACCTGTCTGTGTCCAAAGCCCG 1660
DB 193 GCGCTTACCCCGAGTGTGGCCCACTGCAATGACCCAGGACCTGTCTGTGTCCAAAGCCCG 252

QY 1661 CCAGAGTGGAGGCGCTGTCTGGAGAGTGTGTACGCGCAAGGGCGGTGGCCACCG 1720
DB 253 CCAGAGTGGAGGCGCTGTCTGGAGAGTGTGTACGCGCAAGGGCGGTGGCCACCG 312

QY 1721 CCAGAGGAGGAGATCAAGTGTCCGTTGCGGAGAGATCCCGGCGACCTGCGCGCG 1780
DB 313 CCAGAGGAGGAGATCAAGTGTCCGTTGCGGAGAGATCCCGGCGACCTGCGCGCG 372

QY 1781 TGTCTTACGCCGCCCAACACCTGAAGCCCGTGTCCGCGCTCCGTTGGAGGCGCAAGCGCGCG 1840
DB 373 TGTCTTACGCCGCCCAACACCTGAAGCCCGTGTCCGCGCTCCGTTGGAGGCGCAAGCGCGCG 432

QY 1841 CCAGCGCCAGAGTGGGACACCGCCCGCCATGGCGCGTGGCGCGACACACCCCT 1900
DB 433 CCAGCGCCAGAGTGGGACACCGCCCGCCATGGCGCGTGGCGCGACACACCCCT 492

QY 1901 CGGGCCCTCGCGCGCGCGCGCCACCGCCCAAGGTGACCACTACAAGCCCGCGCTGCGCG 1960
DB 493 CGGGCCCTCGCGCGCGCGCGCCACCGCCCAAGGTGACCACTACAAGCCCGCGCTGCGCG 552

QY 1961 CCACCGCCAGACCGCGCTGCGCTCAAGCGCAACGCGACGCGGTGAGGCGCTCCACCACTCGA 2020
DB 553 CCACCGCCAGACCGCGCTGCGCTCAAGCGCAACGCGGTGAGGCGCTCCACCACTCGA 612

QY 2021 CCTCGGAGACCGCGCTGCGCTCAAGCGCAACGCGACGCGGTGAGGCGCTCCAGACCT 2080
DB 613 CCTCGGAGACCGCGCTGCGCTCAAGCGCAACGCGACGCGGTGAGGCGCTCCAGACCT 672

QY 2081 CGGTGCAAGCGCGCTGCTCGCGCGCCACCGCCCAAGTCCGCG 2124
|||||

DB 673 CGCGTCCAAAGCCCTGGTCTCCGCGCCACCGCCAAAGTCCGCC 716

RESULT 2
AA63355
ID AAX63355 standard; cDNA; 2267 BP.
XX
XX AC
XX AAX63355;
XX
XX 16-JUL-1999 (first entry)
XX
XX Granule bound starch synthase encoding cDNA.
XX
XX Maize; corn; Zea mays; delta-9 desaturase; GBSS; target; substrate;
KW granule bound starch synthase; hammerhead ribozyme; hairpin ribozyme;
KW modulation; gene expression; transgenic plant; cleavage; canola plant;
KW caffeine synthesis; coffee plant; nicotine production; tobacco;
KW fruit ripening; flower pigmentation; lignin production; ss.
XX
XX Zea mays.
XX
XX WO9710328-A2.
XX
XX 20-MAR-1997.
XX
XX 12-JUL-1996; 96WO-US11689.
XX
XX 13-JUL-1995; 95US-0001135.
XX
XX (DOWC) DOWELANCO.
PA
PA (RIBO-) RIBOZYME PHARM INC.
XX
XX Edington BE, Folkerts O, Guo L, McSwiggan JA, Merlo DJ;
PI Merlo PAO, Skokot TA, Young SA, Zwick MG;
XX
XX WPI; 1997-202224/18.
XX
XX Ribozyme which modulates plant gene expression - preferably
PT modulates expression of DELTA-9 desaturase or granule bound starch
PT synthase in maize or canola
XX
XX Example 9; Page 31-33; 155pp; English.
XX
XX The present invention describes an enzymatic nucleic acid molecule (I)
CC with RNA cleaving activity, which modulates the expression of a plant
CC gene. Also described is a gene comprising a cDNA sequence encoding maize
CC Delta-9 desaturase. (I) can be used to modulate expression of a gene,
CC preferably Delta-9 desaturase or a granule bound starch synthase (GBSS)
CC gene, in a plant (preferably a maize or canola plant). (I) can be used
CC to modulate caffeine synthesis in a coffee plant, nicotine production in
CC a tobacco plant, fruit ripening processes in an apple, tomato, pear,
CC plum or peach plant, flower pigmentation in a rose, petunia,
CC chrysanthemum or marigold plant or lignin production in a tobacco,
CC aspen, poplar or pine plant.
XX
SQ Sequence 2267 BP; 465 A; 699 C; 733 G; 370 T; 0 other;

Query Match 22.5%; Score 478; DB 18; Length 2267;
Best Local Similarity 59.8%; Pred. No. 4.4e-61;
Matches 986; Conservative 0; Mismatches 565; Indels 99; Gaps 7;

QY 161 GTGCCACCTTGGCGCTGGACATCGTGTGCTGTGAGTCCGCGCTGGTCCCAAGA 220
DB 403 GCGCCACGCGCGGATGAAGCTGTCTGTGCGCGCGGAGATGGCGCGTGGAGCAAGA 462

QY 221 CGGGCGCGCTGGGCGGTGTGACTGTGCTGCTTATTTAGCTGTGTAAGCGCGGCCACC 280
DB 463 CGGGCGCGCTGGGCGGTGTGACTGTGCTGCTTATTTAGCTGTGTAAGCGCGGCCACC 522

QY 281 GCGTCATGACATTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 340
DB 523 GTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 582


```
Db 1544 CGACTGGTGGACTTTGACACTGTGAAGAAGGCTACTACTGGATCCATATGGAGCCT 1603
QY 1520 TGA-----ACCCGCAAGCTGGACGAGGCTGACGGCGCCCTGGCCGCCACCCGTC 1573
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1604 TCAATGTTGAATGCGATGTTTTCACCCAGCTGATGTGCTTAAGATAGTAACAACAGTTG 1663
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1574 GCCGTGCCAGGAGGTGTTTCGGGGCGCCGCTACCCCGAGATGGTGGCCCAACTGCATCA 1633
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1664 CTAGAGCTCTTGAGTCTATGGCACCTCGCATTTGCTGAGATGATAAAAAATTGCATGT 1723
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1634 GCCAGGACCTGTCTCTGTTCCAAAGCCGCCAGAAAGTGGGAGGCGCTGCTGGAGG 1687
      |||| ||| |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1724 CAGAGGAAGTCTCTCTGGAAGAACCTGCCAAGAAATGGGACACATGCTATTGG 1777
      |||| ||| |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
AAC32824
ID AAC32824 standard; DNA: 2115 BP.
XX
AC AAC32824;
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 779.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
```


Db 1384 AGCAGAAGGGTCTTCATATCTAGTGGAGGCTATTTCCAAAGTTCATGG---GGTCAATG 1440
 QY 1235 TCGAGATGCCATCTGGTACGGCAAGCGCGCTACGAGAAGCTGGTGAACGCCATCG 1294
 Db 1441 TCGAGATGGTTATCTTGGGACTGGAAAGAGAAAGATGGAGGCTCAGATTCTTGAACATG 1500
 QY 1295 GCACCAAGTACAAGGGCGGCCAAGGCGGTGGTCAAGTTCTCGGCGCCCTGGCGGACA 1354
 Db 1501 AAGAGAAGTCCAGGAGGCGGTGGAGTGGCGAAATCAACGTGCCATTGGCTCAT 1560
 QY 1355 TGCTACCGCGCGCGGCTCATGTGTGTCCTCGCTTCGAGCCCTCGAGCCCTGCGGCCGA 1414
 Db 1561 TGATCACTGTGGAGTGCATTCATGTCCCAAGCAGGTTGAGCGGTGGTCTCA 1620
 QY 1415 TCCAGCTGCAGGCCATGCACTACGTTACCGTGGCCGCTGGTAGCCTCCACCGCGCCCTGG 1474
 Db 1621 TTCAGCTGCAGCGAATGAGATATGAACCGCTCCCTATGTGGCATCTACTGGTGACTG 1680
 QY 1475 TCGACACCGTCAAGGAGGCGCTACCGGCTTCCACATGGG-----CGCCCTGACCCCG 1528
 Db 1681 TGGACACTGTAAGATGGCTACACAGTTTCCACATGGAGATTCACAGTCAAGTGTG 1740
 QY 1529 ACAAGCTGGAGAGGCTGACGCCGACCGCTTGGCGCCGCTGCGCGTGCAGCGAGG 1588
 Db 1741 AAGTTGGATCCAGATGATGTATAGCAACAGCAAGAGGCTGTGACAAAGCCGTTGCG 1800
 QY 1589 TGTGTTGGGGCGCGCTACCCGAGATGGTGGCCCAACTGCATCAGCCAGGACCTGTCT 1648
 Db 1801 TATATGAACATCCGAATGCAAGAAATGGTCAAGAACTGCATGACCAAGACTTCTCT 1860
 QY 1649 GGTCCAAAGCCCGCCAGAAAGTGGGAG 1674
 Db 1861 GGAAGGACCTGCGAGGTGTGGGAG 1886

RESULT 7

ABK53215

ID ABK53215 standard; cdna; 2183 BP.

AC ABK53215;

XX

DT 14-AUG-2002 (first entry)

DE Canna edulis granule-bound starch synthase I (GBSSI) cdna sequence.

XX Starch; amylose; cereal crop; commodity starch; thickened foodstuff;

KW bread; baking; pasta; potato; granule-bound starch synthase I; GBSSI; ss.

XX Canna edulis.

OS

XX W0200218606-AL.

PN

XX

PD 07-MAR-2002.

XX

PF 28-AUG-2000; 2000WO-US23494.

XX

PR 28-AUG-2000; 2000WO-US23494.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Broglie K, Lightner J;

XX

XX WPI; 2002-404556/43.

DR

XX

XX

PT Novel starch isolated from cereal crop with a modification of amylose

XX content and/or amylose weight distribution relative to commodity starch

PT from same species of the cereal crop, for preparing thickened

XX foodstuffs

PS

XX Example 1; Page 37-38; 44pp; English.

XX

CC The present invention relates to a new starch comprising amylose,

XX where the starch is isolated from a cereal crop, and comprises a 50%

CC

CC increase in average molecular weight relative to commodity starch from
 CC the same species of the cereal crop from which starch was isolated.
 CC The starch of the invention is useful for preparing thickened foodstuffs
 CC by combining the starch, a foodstuff and water and cooking the invention
 CC composition as necessary to produce a thickened foodstuff. The invention
 CC is also useful for preparing foodstuffs such as breads, baked goods,
 CC pastas, etc. The present nucleic acid sequence represents the Canna
 CC edulis granule-bound starch synthase I (GBSSI) cdna sequence in clone
 CC ectlc.pk007.015 that was used in the methods of the invention for
 CC generation of a full-length sense Canna edulis GBSSI construct.
 XX

SQ Sequence 2183 BP; 636 A; 428 C; 516 G; 603 T; 0 other;

Query Match 12.1%; Score 258; DB 24; Length 2183;

Best Local Similarity 51.6%; Pred. No. 3.4e-29;

Matches 794; Conservative 0; Mismatches 650; Indels 96; Gaps 5;

QY 176 TGGACATCGTGATGTTGCTGTGAGTGGCCCTTGGTCCCAAGACGGCGGCTGGGCG 235

Db 357 TGAACCTTGGGTTTGTGGTTGTAGTAGCTCCATGGAGCAAAAACCTGGGGCCTTGGCG 416

QY 236 ATGTGACTGGTGGCTGCTTATGAGCTGTTCAAGCGCGGCCACCGCTCATGACCATG 295

Db 417 ATGTTCTTATAGGATGCCACTGCTATGGCTGCAATTTGGGCACAGGTCATGACCGTGG 476

QY 296 CCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGTCGTGGACAT----- 350

Db 477 CGCCACGATATGACCAATATAAGATATCTGGATACAAAGTCTCCACTTGGTTAAAG 536

QY 351 -----CATGGCGAGAGGTCGCTACTTCCACTCATCAAGAGGCGTGCACCGG 403

Db 537 TTGGGATAGATTGAAACTGTCGCTCTTCCACTGTACAAAAGGGAGTTGATCGG 596

QY 404 TGTGATTGACCAACCCCTGTTCTGTCGCAAGCTGTGGCAAGACCGGCTCCAGCTGT 463

Db 597 TTTTGTGGATCACCCTATGTTCTCGAAGAGTTTGGGGAACAGAGGAAATAT 656

QY 464 AGGCCCCCGCTCGGCGCTGACTACTGAGCAACCAAGCGTTCGCCCTGTTCTGCA 523

Db 657 ATGGTCTGTTACAGAACAGATTATGACAGAACTCACTAAGATTGACGCTTTTGTGCC 716

QY 524 AGGCGCTATGAGCTGCCCGCTGCTGCTGCGCCCTTGGGCGCGGAGGAGTGC----- 576

Db 717 TGGCAGCTCTGGAAGCTCCAAAGACTTCTAAATCTCAACACAGCAAAATACTATTCTGGAC 776

QY 577 -----GTCTTCTGTCGCAAGCTGCTGCTGCGCCCTTGGGCGCGGAGGAGTGC----- 576

Db 777 CATATGAGATGATGTGTGTTTATGCCAAGATGGCATCTGCTCTACTGCCCTGCT 836

QY 620 TGCTGAAGGAGGATACAGCCCAAGGCGCCAGTTTCACCAAGGCCAAAGTCCGTTGCTGCTA 679

Db 837 ACTTGAAGAACTATGTACCAATACATGTTATTTACATGAATGCTAAGTTGCAATTTGCA 896

QY 680 TCCACACATCGCTTCCAGGCGCGCATGTGGGAGGAGGCTTTCAAGGACACAGAGTGC 739

Db 897 TTCATAATATGCTTACCAGGCGCGATTGCTTTTCGAGCTTTTGAACCTCTCTAATCTCC 956

QY 740 CCCAGGCGCGCTTTGACAAGCTGGCCCTTCTGCGAGGCTATGCCAAGSTTTTACACTGAGG 799

Db 957 CCATAAATTTAATCTTCAATTTGATTTGATGATGATGATGACAAACCT----- 1006

QY 800 CCACCCCATGGAGGAGGACGAGAACGCCCGCTGACGGGAAGAACCTACAGAAGATCA 859

Db 1007 -----GTGAAGAGGAAGAAATAA 1025

QY 860 ACTGGCTGAAGGTTGGCTATTCGCCGCCACAGAGCTGGTGTGCTGCTGCCCAACTACG 919

Db 1026 ATTGGATGAAGGCTTGAATAATAGATGTAGGTTGCTTACCGGTGAGCCCATATATG 1085

QY 920 CGACCGAGATCGCTGCCGATGCCCGCGGTGTGGAGCTGGACACGCTCATCCGCCCA 979

Db 1086 CCCAAGAGCTTGCTCAGGGGTAGAGAAGGTTGTGAGTTGGGCAATATCTCGCATGA 1145

QY 980 AGGCAATGAGGCAATGTGAACGGCATGGACATTTGAGGAGTGGAAACCCCAAGACGACA 1039
Db 1146 ARACCAATCTGGAATAGTAAATGGATGGACACACGAGTGGATCCATTAACAGACA 1205
QY 1040 AGTTCCTGTCTGGCCCTAGCAGCAGACAGCGCTCTACGCCGCAAGCGCGCCCAAG 1099
Db 1206 AATATATTCTACAAACTACGATGCAACAACCTATTGGATGCAAAACCTCTCTGAAG 1265
QY 1100 AGCCCTGAGCGAGTGGGCTGCTGTGGAACCCACCGCCCTGTGCGCCTTCA 1159
Db 1266 AAGCTTTGAAGCTGAGTGGGCTGCTGTGGAACCAAAACAAAGCTGTTTGGCCTTG 1325
QY 1160 TCGCCGCTGGAGGAGCAAGAGGTGTGACATCATCTCTGGCCGCTGCCCAAGATCC 1219
Db 1326 TTGAAGACTAGATGAGCAAGAGCTACACATCTAGTGCAGCAATTCAGAACTTC 1385
QY 1220 TGGCCACCCCAAGTGTGAGATGCCATCTCTGGGTACCGGCAAGCGCCCTACGAGAGC 1279
Db 1386 T--TTGTGAGAAATGTTCAAGTGATAGTACTTGGCACTGGCAAGAGAAAGTTGGAGAGTG 1442
QY 1280 TGGTGAACGCCATCGGCACCAAGTACAAAGGCGCGGCAAGGCGTGTGTCAGTTCTCGG 1339
Db 1443 AACTTACATTTAGGAAATGTTCCAGACAAATTCAGACACATCTCAAAATCAAG 1502
QY 1340 CGCCCTGGCGCACATGCTCACCGCGCGCGGCGGACTTCATCTGTGTCGCGCTCG 1399
Db 1503 TTCTTTAGCTCATGCAATCATGCGAGGAGCTGATATCTTGTATTCAGAGCAGATCG 1562
QY 1400 AGCCGTGGCGCTGATCCAGCTGCAGCCATGACATACGATACGTCGCGGTGAGCCT 1459
Db 1563 AACCCTGTGGCCCTCATTTCAGCTTCAGGCCATGGGATATGGAATCTCCCTATGTATGASCA 1622
QY 1460 CCACGCGCGCTGTGCGACACGCTCAAGAGGCGCTCACCGGCTTCCACATGGCGCC- 1518
Db 1623 CCACTGTGACCTGTGACACCTGTCAAGAGGCTTCACTGGCTTCCATATGAGGCCCT 1682
QY 1519 -----CTGACCCCGCAAGCTGGAGAGCTGAAGCGGCTGAGCGGCGCCCTGGCGCCACCGTGC 1573
Db 1683 TCAGTGTGGAGTGTGATGCGGTAGACAAAGCTGATGTACAAAGATTGTGAAACACGA 1742
QY 1574 GCCGTGCCAGGAGTGTGTTGGGGCGGCGCTACCCCGAGATGTTGGCCAACTGCATCA 1633
Db 1743 AAAGGGCCCTCAAGTCTATGGAACACCTGCTTTTGGAGATGATCAAGAACTGCATGA 1802
QY 1634 GCCAGGACCTGCTGCTGTCCAAAGCCCGCCAGAGTGGGA 1673
Db 1803 ACCAAGATCTCATGGAAGGACCTGCAAGAGTGGGA 1842

RESULT 8
AAC86435
ID AAC86435 standard; cdna; 2807 BP.
XX AAC86435;
AC AAC86435;
XX 01-MAR-2001 (first entry)
DT Wheat starch synthase clone wSSIID cdna.
DE Wheat starch synthase clone wSSIID cdna.
XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
KW food product; adhesive; ss.
XX Triticum aestivum.
XX W0200066745-A1.
XX 09-NOV-2000.
XX 28-APR-2000; 2000WO-AU00385.
XX 29-APR-1999; 99AU-0000052.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
FA

PA (GOOD-) GOODMAN FIELDER LTD.
XX (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
PI Morell M, Li Z, Rahman S, Appels R;
XX WPI; 2000-647602/62.
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
PT WST-II, useful in modifying plant starch content and/or composition -
XX Disclosure; Fig 2; 211pp; English.
XX The present invention relates to novel protein and coding sequences from
CC wheat. The proteins are wheat starch synthases, designated SSII and
CC SSIII. These can be used in the modification of plant starch content or
CC composition, and to screen plants to identify mutations which affect
CC starch content and composition. The starch can then be used in food
CC products, such as flour, and in films, coatings, adhesives, building
CC materials and packaging materials.
XX Sequence 2807 BP; 549 A; 844 C; 907 G; 507 T; 0 other;
SQ
Query Match 10.4%; Score 221; DB 21; Length 2807;
Best Local Similarity 56.2%; Pred. No. 7.6e-24;
Matches 485; Conservative 0; Mismatches 360; Indels 18; Gaps 3;
QY 851 AGAAGATCAACTGCTGAAGGGTGGCATATCGCGCGCGCAAGCTGTGCTGTGCG 910
Db 1620 AGCAGCCCACTACTTTCGCGCGCGCTGAGATGGCGGACAGTTGCTGCTGTGAGCC 1679
QY 911 CCAACTACGGCAACAGATCGCTCCGATGCCCGCGGGTGTGGAGTGGACACCGTCA 970
Db 1680 CCGGTTACCTGTGGAGCTCAAGCGGTGGAGGCGGTGGGGCTTTCACGACATCATAC 1739
QY 971 TCC---CGCCCAAGGCGATTGAGGCGCATTTGAACGGCATGACATTCAGGAGTGGAAAC 1027
Db 1740 GGCAGAGCATGGAAGACCCCGCGCATTCGCAAGCGCATGCAACAATGAGTGGAAAC 1799
QY 1028 CCAAGACCGACAAGTCTCTGTCTGCGCCTACGACCAAGAACAGCGTCT----- 1075
Db 1800 CCGAGGTGGAGCTCCACCTCAAGTCGGACGCTACACCACTTCTCCCTGGGAGCGTGG 1859
QY 1076 ACGCCGCAAGCGCGCCCAAGGAGGCGCTGAGGCGGAGCTGGCTGCTGTGAGCC 1135
Db 1860 ACTCCGCAAGCGCGCTGCAAGGAGGCGCTGAGGCGGAGCTGGCTGCTGTGAGCCG 1919
QY 1136 CCACGCGCCCTGTCTGCGCTTCATCGCGCGCTTGGAGGAGCAGAAGGCTGTGGACATCA 1195
Db 1920 CCGAGTGGCGCTGCTCGGCTTCATCGCGCGCTTGGAGGAGCAGAAGGCGTGGAGATCA 1979
QY 1196 TCCTGGCGCGCTGCGCCCAAGATCTCTGCGCACCCCGAGTGCAGATGCCATCTGGTA 1255
Db 1980 TCGCGGAGCGCATGCCCTGGATCGTG---AGCCAGGAGCTGCTGCTGCTGCTGCGCA 2036
QY 1256 CCGCAAGCGCGCTTACGAGAGCTGTTGAAGCGCATCGCACCAAGTACAGGCGCGCG 1315
Db 2037 CCGCGCGCGCTGAGAGCATGCTGCGGCACTTCAGCGGAGGAGCACCACAGCAGG 2096
QY 1316 CCAAGGCGCTGCTCAAGTCTCGCGCGCTTGGGCGCATGCTACCGCGCGCGCGGACT 1375
Db 2097 TCGCGGCTGGTGGGCTTCTCCGTGCGCTTGGCGCACCGGATCACGCGCGCGCGGACG 2156
QY 1376 TCATGCTGCTGCGCTGCGCTTTCGAGCGCTTGGGCTGATCAGCTGACGCCATGCACT 1435
Db 2157 CGCTCTCATGCCCTCCCGGTTTCGAGCGCTGCGGCTTGAACCGAGCTTTACGCGCT 2216
QY 1436 ACCTGACCGCTGCGCTGCTGCTTCCAGCGCGCTTCCAGACCGCTCAAGGAGGCGG 1495
Db 2217 ACAGCAGCTGCGCTGCTGCTGCGCGCTTGGGAGGAGCAGCGCTGCGCGCGCTGCG 2276
QY 1496 TCACCGCTTCCATGAGGCGCGCTTGAACCGCGCAGCAAGCTGAGCGCTGACCGCGAG 1555
Db 2277 ACCCTTCAACCACTCCGCGCTTGGAGTGGACGCTTTCGCGCGCGCGCGGAGCGCAAGCTGA 2336

QY 1556 CCCTGGCCGACACCGTGGCGGTGCGAGAGGTGTTTGGGGCGCGCTACCCGAGA 1615
 Db 2337 TCGAGGCGCTGGGCACTGCGCTCCGCACTTACCGGACTACAAGGAGAGCTGGAGGGCC 2396
 QY 1616 TGGTGGCAACTGCATCAGCAGGACCTGTCTGTCTCCAAAGCCGCCAGAGTGGGAGG 1675
 Db 2397 TCCAGGAGCGGGCATGTCGAGGACTTCAGCTGGGAGCATGCCGCCAAGCTCTACGAGG 2456
 QY 1676 GCCTGCTGGAGAGGTGTGTAC 1698
 Db 2457 ACGTCTCTCTCAAGGCCAAGTAC 2479

RESULT 9

AAC86411

ID ARC86411 standard; cDNA; 2842 BP.

XX AAC86411;

AC AAC86411;

XX 01-MAR-2001 (first entry)

DT Wheat starch synthase II coding sequence SEQ ID NO: 3.

DE Wheat;

DE starch synthase; SSII; SSIII; starch content; starch synthesis;

KW food product; adhesive; ss.

KW Triticum aestivum.

OS WO200066745-A1.

XX 09-NOV-2000.

XX 28-APR-2000; 2000WO-AU00385.

XX 29-APR-1999; 99AU-0000052.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (GOOD-) GOODMAN FIELDER LTD.

PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

XX Morell M, Li Z, Rahman S, Appels R;

PI WPI; 2000-647602/62.

XX P-PSDB; AAB37567.

DR Nucleic acid molecules encoding wheat starch synthase (WST)-I and

PT WST-II, useful in modifying plant starch content and/or composition -

XX Claim 1; Page 155-158; 21pp; English.

PS The present invention relates to novel protein and coding sequences from

XX wheat. The proteins are wheat starch synthases, designated SSII and

CC SSIII. These can be used in the modification of plant starch content or

CC composition, and to screen plants to identify mutations which affect

CC starch content and composition. The starch can then be used in food

CC products, such as flour, and in films, coatings, adhesives, building

CC materials and packaging materials.

XX Sequence 2842 BP; 561 A; 851 C; 919 G; 511 T; 0 other;

SQ Query Match

CC Best Local Similarity 10.4%; Score 221; DB 21; Length 2842;

CC DE Local Similarity 56.2%; Pred. No. 7.6e-24;

CC Matches 485; Conservative 0; Mismatches 360; Indels 18; Gaps 3;

QY 851 AGAAGATCAACTGGCTCAAGGGTGGGATTATCGCCCGCAAGCTGGTGACTGTGCGC 910

Db 1620 AGCAGCCAACTACTTCGCGCGCGCTGAGATGCGGACCAAGTTGCTGTGTAGGCC 1679

QY 911 CCAACTACGGACCGAGATGCTGCCGATGCCCGCGCGGTGTGGAGCTGGACACCGTCA 970

Db 1680 CCGGGTACCTGTGGGAGCTCAAGACGCTGGGGGCGCTTCACGACATCATAC 1739

QY 971 TCC---GGCCCAAGGGCATTTGAGGGCATTTGAACGGCATTGACATTTGAGGAGTGAACC 1027
 Db 1740 GGCAGAACACTGTGAAGACCCGCGCATGTTCAACGGCATTCGACAACTGAGTGAACC 1799
 QY 1028 CCAAGACCCACAAGTTCTGTCTGCGCCCTAGCACAGAACAGCGTCT-----1075
 Db 1800 CCGAGTGGAGCTCCACCTCAAGTCTGAGCGGTACACCAACTTCTCCCTGGGAGCGTGG 1859
 QY 1076 AGCCGCGCAAGCCCGCCGAAGGAGGCGCTGCAGCGAGCTGGGCGCTGCTGTGGGACC 1135
 Db 1860 ACTCGGCAAGCGGCAAGTGCAGAGAGGCGCTGCAGCGAGCTGGGCGCTGAGGTCGCG 1919
 QY 1136 CCACGCGCCCGCTGTTCCGCTTCATCGCGCGCTTGAAGAGCAGAGGTTGTGGACATCA 1195
 Db 1920 CCGAGTGGCGCTGCTCGCTTCATCGCGCGCTTGAAGAGGCGGTGGAGATCA 1979
 QY 1196 TCCTGGCGCGCTGCCAAGATCCTGGCCACCCCAAGTGCAGATCGCCATCTCTGGGTA 1255
 Db 1980 TCGGAGCGCATGCCCTGGATCTG---AGCCAGGAGCTGCAGCTGTCTATCTGGGCA 2036
 QY 1256 CCGCAAGCGCGCTTACGAGAGCTGTGTGAACCGCATCGGCACCAAGTACAGGCGCGCG 1315
 Db 2037 CCGCGCGCACGACTGGAGAGCTGTGCGGCACTTCGCGGAGGACCAACGACAGG 2096
 QY 1316 CAAGGGCGTGTCAAGTTCCTGGGCGCGCTTGAAGAGCAGATGCTACCGCGCGCGGCT 1375
 Db 2097 TCGCGGCGTGGTGGGTTCTCCGCTGCGCTGGCGACCGGATCACGCGCGCGCGGCG 2156
 QY 1376 TCATGCTGTGCTCGCGCTTCGAGCGCTTGGGCGCTTGGGCGCTTACAGCTGCACGCT 1435
 Db 2157 CGTCTCTATGCTTCCCGGTTGAGCGCTTGGGCGCTTGAACAGCTTACCGCATGGCT 2216
 QY 1436 ACGTACCGTCCCGTGTGAGCTTCCACCGCGCGCTTGGTGCAGACCGCTCAAGGAGGCG 1495
 Db 2217 ACGCACCGTCCCGTGTGTCGACCGCTGCGGCGGTGAGGAGACCGCTGCGCGGCTCG 2276
 QY 1496 TCACGCGTTCACATGGGCGCTTGAACCGCGACAACTGGAGAGGCTGACCGCGGAG 1555
 Db 2277 ACCCTTCAACCACTCCCGCTCGGCTGAGGCTTGCAGCGCGCGGAGGCGACAGCTGA 2336
 QY 1556 CCCTGGCGCGCACCGTGGCGCTGCCAGGAGGTGTTTGGCGGCGCGCTACCCCGAGA 1615
 Db 2337 TCGAGGCGCTCGGCACTGCTCCGCACTACCGGAGTACAAGGAGAGCTGGAGGGGCC 2396
 QY 1616 TGGTGGCAACTGCATCAGCAGGACCTGTCTGTCTCAAGCCCGCGCAGAAAGTGGGAGG 1675
 Db 2397 TCCAGGAGCGGCGCATGTCGAGGAGCTTCAGCTGGGAGCTGCCCGCAAGCTCTACGAG 2456
 QY 1676 GCCTGCTGGAGGAGTGTGTAC 1698
 Db 2457 ACGTCTCTCTCAAGGCCAAGTAC 2479

RESULT 10

AAV29755

ID AAV29755 standard; DNA; 2097 BP.

XX AAC29755;

XX 11-SEP-1998 (first entry)

DE Zea mays soluble starch synthase IIB gene.

KW SBR; starch-encapsulating region; fusion vector;

KW soluble starch synthase IIB; glucosyl transferase; ss.

OS Zea mays.

XX Key

FH Location/Qualifiers

FT 1.2097

FT /tag= a

FT /product= soluble starch synthase IIB

XX

PN W09814601-A1.
XX 09-APR-1998.
XX 30-SEP-1997; 97WO-US17555.
XX 30-SEP-1996; 96US-0026855.
XX (EXSE-) EXSEED GENETICS LLC.
XX Guan H, Keeling P;
XX WPI: 1998-240100/21.
XX P-PSDB; AAW56487.
XX Hybrid polypeptide comprising starch-encapsulating region and
XX protein - useful for, e.g. producing protein(s) resistant to
XX degradation by stomach acids
XX Example 2; Page 36-38; 156pp; English.
XX The sequence is that of the soluble starch synthase IIb gene.
XX It can be used in the production of a hybrid polypeptide
XX comprising a starch-encapsulating region (SER) fused
XX to a payload protein. The hybrid polypeptide can be used to make
XX modified starches comprising the payload protein, selected from,
XX e.g. hormones, growth factors, antibodies, enzymes, dyes,
XX immunoglobulins, etc. The modified starch can also be used
XX to provide grain feeds enriched in amino acids. By encapsulating
XX the payload protein in starch, it is more resistant to
XX degradation by stomach acids.
XX Sequence 2097 BP; 409 A; 575 C; 696 G; 417 T; 0 other;
XX
XX Query Match 10.3%; Score 219.2; DB 19; Length 2097;
XX Best Local Similarity 50.8%; Pred. No. 1.4e-23;
XX Matches 801; Conservative 0; Mismatches 673; Indels 108; Gaps 7;
XX
QY 144 TGTTACTGGTGCACCTGCTGGCTGGACATGCTGATGTTGCTGAGGT 203
Db TGAACCTGGCCCTTGGCTGGCCCTAATGTGATGAACGTGCTGCTGGCTTCTGAATG 647
QY 204 CGCCCTTGGTCCAAAGCGGCGCTGGGGATGCTGCTGGCTGCTGCTGCTGCTGCT 263
Db TGCTCTCTTCTGCAAGACAGGTGGCTTGGAGATGCTGCTGGCTGCTGCTGCTGCTGCT 707
QY 264 GGTCAAGCGGCGCACCGCTGATGACCTTGGCCCTGCTGACCAAGTACGCTGACGC 323
Db GGCAGAGGAGGACCGCTGTTAATGCTGATACCAAGATATGAGAGTATGCCGAGC 767
QY 324 CTGGGACACCTCGGTGCTGGACATCATGGCGGAGAAAGTCCGCTACTTCCACTCCAT 383
Db CCGSG--ATTAGTGTAAAGGAGAGCTTACAAGGTAGCTGACAGAGTATCAGAAGTTACT 825
QY 384 CAAGAAGGGGTGACCGGTGGATGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTG 443
Db TAITTTCACTCTTACATTTGATGAGTGTGATTTTCTATTTAGTAAAGCCCTCCCTTCGG 885
QY 444 CAAGACCGGCTCCAAAGTGTACGCGCCCGCTCGGCGCTGACTACCTGGACCAACCAA 503
Db 886 CAC-----CGGCACAAATATTTATGCGGGGAGAAAGATTTGATATTTGAA 932
QY 504 GCGCTCGCCCTGTGTCAGGCGCGCTGATGAGCTGCGCGGCTGCTGCTGCTGCTGCTG 563
Db GCGCATGATTTTGTCTCAAGGCGCGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTG 992
QY 564 -----CGGCGAGGACTGGCTTCTGCTGGCAACACTGGGCTGCTGCTGCTGCTGCT 611
Db 993 TACTGTCTATGATGAGCACTTAGTGTTCATGCTAATGATGGCATACCGCACTTCT 1052
QY 612 GCCGCTCGCTGAGGACGAGTACCAAGCCCAAGGCGGCTGCTGCTGCTGCTGCTGCTG 671
Db 1053 GCCTGTCTATCAAGGCTTATACCGGGACAAATGTTGATGACGATGCTGCTGCTGCT 1112

QY 672 GCTGGCTATCCAAACATCCCTTCCAGGCGCGCATGTGGAGGAGGCTTTCAGGACAC 731
Db 1113 GCTTGTATACACACATTCCTCATCAGGCTGCTGGCCCTGTAGACGACTTCGTAATTT 1172
QY 732 GAAGCTGCCCCAGGCGCCCTTTGACAAAGCTGGCCCTTTCGACGGGCTATCCAAAGGTTTA 791
Db 1173 TGACTTGCCTGAACACTACATCAGCCACTTCAAACCTATATGACAAACATTTGGTGGGATCA 1232
QY 792 CACTGAGGCCACCCCATGAGGAGGACGAGAGCCGCCCTGACGGGAAAGACCTACAA 851
Db 1233 C-----1233
QY 852 GAAGATCACTGGCTGAAGGTGGCAATTTGCGGCGCGACAAAGCTGTGTGCTGTGCTGCC 911
Db 1234 ----AGCAACGTTTTTCTGCTGCGGGCTGAAGACGGCAGACCGGCTGTGACCGCTTAGCAA 1289
QY 912 CAACTAGCGACCGAGATCGCTGCCGATGCGCGCGGCTGTGAGCTGGACACCGCTCAT 971
Db 1290 TGGCTACATGTGGGAGCTGAAGACTTCGGAAGCGGCTGGGGCTCCACGACATCATAA 1349
QY 972 CC---GCGCCAAAGGCGCATTTGAGGCGCATTTGAACGCGCATGGACATTTGAGGAGTGAACCC 1028
Db 1350 CCAGAAGCACTGGAGCTGAGGCGCATCTGTGAACGCGCATCGACATGAGCGAGTGAACCC 1409
QY 1029 CAAGACGCAAGTTCCTGCTGCGCCCTACGACCGACGACAGC-----GTCTA 1076
Db 1410 CGCTGTGACGTGCACCTCCACTCCGACGACTACACAACTATACAGCTTCGAGACGCTGGA 1469
QY 1077 GCGCGGCAAGGCCCGCCAGGAGGCGCTGACGCGGAGCTGGCGCTGCTGCTGCTGCTG 1136
Db 1470 CACCGCAAGCGGAGTGAAGGCCCTTCGACGGCGCTGGCGCTGCGAGTCCCGGGA 1529
QY 1137 CACGCGCCCTGCTTCGCTTCATCGCGCGCTGAGGAGGAGGAGGCTGTGGACATCAT 1196
Db 1530 CGAGCTGCCACTGATGGGTTTCATCGGCGGCTGGACCAACAGGAGGCGCTGGACATCAT 1589
QY 1197 CTTGGCGGCTGCGCCAAAGATCTTGGCCACCCCAAGTGCAGATCGCCATCTCGGGTAC 1256
Db 1590 GCGCGAGCGGATCCTGATGATC---GCGGGGACGAGCTGAGCTGCTGCTGCTGCTG 1646
QY 1257 CCGCAAGCGGCTTACGAGAGCTGTTGAACGCGCTCGGACCAAGTACAAGGCGCGCGC 1316
Db 1647 CCGCGGCGCCACCTGAGGAGCATGCTGCGGGGCTTCAGTTCGNGACACGACAAAGT 1706
QY 1317 CAAGGCGGTGCTAAGTTCTCGGCGCCCTTGGCGACATGCTACCGCGCGCGCGGCTT 1376
Db 1707 GCGCGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1766
QY 1377 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1436
Db 1767 CTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1826
QY 1437 CGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1496
Db 1827 CCGGACCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1886
QY 1497 CACCGGCTTCCACATGCGGCGCCCTGAAACCCCGACAGCTGAGCTGAGGCTGAGCGGCGC 1556
Db 1887 CCGCTTCAACGACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1946
QY 1557 CTTGGCGCCACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1616
Db 1947 CGAGCGCTTCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2006
QY 1617 GGTGCCCAACTGCTACGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1676
Db 2007 CAGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2066
QY 1677 CTTGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1698
Db 2067 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088

QY 1304 ACAAGGCGCGCCACAGGGGTGTTCAAGTCTCTGGCGCCCTGGCGCACATGCTCACCG 1363
 DB 2158 ACCACACAAAGGTGCGCGGTGGGTCTCTCCGTGCGCTTGGCGCACCGATCACGG 2217
 QY 1364 CCGGCGCGGACTCATGCTGGTGGCTCGCGCTTCGAGCCCTCGCGCTGATCCAGCTGC 1423
 DB 2218 CGGGGGGAGCGGCTCTCATGCTCCCGGTTGAGCGGCTGCGGCTGAACCACTCT 2277
 QY 1424 ACGCCATGACATAGGTACCGTGGTGTAGCTTCACCGCGCGGCTGGTGGACACCG 1483
 DB 2278 ACGCCATGGCTACGGCACCGTCCCGCTGTCACACCGCGCTGCGCGCTCAGGGACACCG 2337
 QY 1484 TCAAGAGGGCGTACCGGCTTCCACATGGCGCGCTGAAACCCCGACAACTGGACAGG 1543
 DB 2338 TGGCGGCTTCGACCCCTTCAACACCTCCGGGCTCGGGTGGAGTTCGACCGCGCGAGG 2397
 QY 1544 CTGACGCGGACGCGCTGGCGCGCACCGTGGCGGTGCGGAGGTTGTTTCGGGGGGCC 1603
 DB 2398 CGCACAAGCTGATCGAGGCGCTCGGCACTGCTCGCACCTACCGAGACTTCAAGGAGA 2457
 QY 1604 GTTACCCGAGATGGTGGCACTGATCAGCCAGGACCTGCTCTGTCCCAAGCCGCC 1663
 DB 2458 GCTGGAGGGCCCTCCAGGACGCGGCGATGTCGACGAGCTTCAGCTGGAGCACGCCCA 2517
 QY 1664 AGAAGTGGGAGGCGCTGCTGGAGGAGGTGGTGTAC 1698
 DB 2518 AGCTTACGAGGAGCTGCTGTCAGGCCAAGTAC 2552

RESULT 12

AAV70958
 ID AAV70958 standard; DNA; 2423 BP.
 AC AAV70958;
 XX 26-AUG-1999 (first entry)
 XX DNA encoding maize starch soluble synthase IIb.
 KW Non-glycogen-like polysaccharide production; fermentation;
 KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;
 KW non-starch branching gene; amylopectin; amylose; plant-like starch;
 KW maize starch soluble synthase IIb; ss.
 XX Zea mays.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1632
 FT FT /**tag= a
 FT FT /note= "this base represents a nucleotide missing from
 FT FT the sequence given in the specification. It
 FT FT is included to maintain the nucleotide
 FT FT numbering given in the specification for this
 FT FT sequence"
 XX
 XX WO9844780-A1.
 XX
 XX 15-OCT-1998.
 XX
 XX 03-APR-1998; 98WO-US06660.
 XX
 XX 04-APR-1997; 97US-0042939.
 XX
 XX (EXSE-) EXSEED GENETICS LLC.
 XX
 XX Guan H, Keeling PL;
 XX WPI; 1998-568285/48.
 XX P-PSDB; AAV70892.
 XX
 XX Producing non-glycogen-like polysaccharides in bacteria, fungi or
 XX plants - transformed with genes for enzymes involved in starch or
 PT

PT glycogen synthesis allows fermentative production of starches with
 engineered properties

PS Disclosure; Fig 50; 150pp; English.

XX The specification describes a method for the production of
 non-glycogen-like polysaccharides in a host. The method comprises
 CC transforming a host, suitable for fermentation, with genes encoding
 CC starch- or glycogen-synthesis enzymes, and fermenting the transformants.
 CC The specification also describes hosts transformed with a gene active
 CC in glycogen synthesis and at least one non-starch branching gene,
 CC involved in production of amylopectin or amylose in its original host.
 CC The method is used to produce plant-like starches by fermentation and
 CC new starches in plants. These starches are useful for all food and
 CC non-food applications of starch. The present sequence is used in
 CC the course of the invention.

XX Sequence 2423 BP; 466 A; 656 C; 804 G; 496 T; 1 other;

Query Match 10.2%; Score 217.2; DB 19; Length 2423;
 Best Local Similarity 50.5%; Pred. No. 2.7e-23;
 Matches 799; Conservative 1; Mismatches 674; Indels 108; Gaps 7;

QY 144 TGTTACTGTGCCACTGTGCCACTTGGCGCTGGACATGCTGTTGCTGCTGAGGT 203
 DB 588 TGAACCTGGCCCTTTGGCTGGCCCTAATGTGATGAACCTGCTGCTGCTTCTGATG 647
 QY 204 CGCCCTTGTGTCGAAGACGGCGGCTTGGCGATGTGACCTGGTGGCTGCTGCTGATG 263
 DB 648 TGCTCCCTTCTGAAGACAGGTGGCTTGGAGATGTGCTGGTCTTGGCTAAAGCTCT 707
 QY 264 GGTCAAGCGGCGCACCGCTCATGACCATTCGCCCTGCTACGACGATGCTGAGCG 323
 DB 708 GCGGAGGAGGACACCGCTGTATGGTGTATACCAAGATATGAGAGTATGCGGAGC 767
 QY 324 CTGGGACACCTCGGTGGTGGACATCATGGGCGAGAGGTCCGCTACTTCCACTCAT 383
 DB 768 CCGGG--ATTAGTGTGAAGGAGCGTTACAAGGTAGCTGGACAGATTCAGAAGTTACT 825
 QY 384 CAAGAAGGGCGTCACCGCGTGGATGACCCCTGGTTCCTGGCCCAAGGCTGCGGG 443
 DB 826 TATTTTCACTCTTACATTGATGGATGATTTTGTATTCGTAGAAGCCCTCCCTCCGG 885
 QY 444 CAAGACCGGCTCCAAGCTGTACGGCCCGCTCGCGCTGCTACTACCTGGACACCAAA 503
 DB 886 CAC-----CGGCACAAATAATTTATGGGGAGAAAGATTGGATATTTGAA 932
 QY 504 GCGCTTCGCCCTGTCTGCAAGGCGGCTATTGAGGCTGCCCGCTGCTGCCCTCGGCC 563
 DB 933 GCGCATGATTTGTCTGCAAGGCGGCTGTGAGGTTCATGTTGCTTCCATGTGGCGG 992
 QY 564 -----CGCGGAGGACTGCTCTGCTGGCCACAGCTGGCCTCGGCTCGGCTGGT 611
 DB 993 TACTGCTATGTTGATGTCACCTTAGTTTTCATGTTGATGATGTCATGCGCACTTCT 1052
 QY 612 GCCCGTCTGCTGAAGGAGGACTACCAAGCCCAAGGSCCAAGTTCAACCAAGCCCAAGT 671
 DB 1053 GCCTGTCTATCAAAGGCGCTATACCGGACAAATGGTTGATGTCAGTATGCTCGCTCT 1112
 QY 672 GCTGCTATCCACAACATCGCCTTCAGGGCGGCGCATGTGGGAGGAGGCTTTCACAGAC 731
 DB 1113 GCTTGTGATACACACATTTGCTATCAGGGTCTGTCGCGCTGTAGAGCTTCTGTAAT 1172
 QY 732 GAAGTGGCCCGAGCGGCTTTGACAAGCTGGCCTTCTCGGACGCGCTATGCCAAGGTTA 791
 DB 1173 TGACTGCTGAACACTACATCGACCTTCAAACTGTATGACAACATTTGGTGGGATCA 1232
 QY 792 CACTGAGGCCACCCCAATGAGGAGGACGAGAGGCCCGCTGACGGGAAAGACCTTACAA 851
 DB 1233 C-----
 QY 852 GAAGATCAACTGGCTGAAGGTGGCATTTATCGCGGCCGACAAAGCTGGTGAAGTGTGCGC 911

Db 1234 ----AGCAACGTTTTTCTGGGGGTGAAGACGGACACCGGGTGTGACCGTTAGCAA 1289
QY 912 CAACCTAGCCGACCGAGATCGTCCGATGCGCGGGGTGTGAGCTGGACACCGTCAT 971
Db 1290 TGGCTACATGTGGGAGTGAAGACTTCGGAAGGCGGTGGGCTCCACGACATCATAAA 1349
QY 972 CC---GCGCCAAAGGCAATGAGGCAATGTGAACGGCATGCACATGAGGAGTGAACCC 1028
Db 1350 CCAGAACGACTGNAAGTGCAGGCAATGTGAACGGCATGCACATGAGGAGTGAACCC 1409
QY 1029 CAGACCGCAAGTCTCTGCTGGCCCTACGACCAAGAACAGC-----GTCTA 1076
Db 1410 CGTGTGGAGCTGCACCTCCACTCCGAGGACTACACCACTACAGCTTCGAGAGCTGGA 1459
QY 1077 CGCGGCAAGCGCCGCCAAGGAGGCCCTGAGGCGGAGCTGGGCTGCTGTGGACCC 1136
Db 1470 CACGGCAAGCGCAGTGCAGGCGCCCTGCAGCGGAGCTGGGCTGCAGGTCCGCGA 1529
QY 1137 CACGCGCCCTCTGCTCCCTTCATCGGCGCCCTGGAGGAGCAGAAAGGTGTGACATCAT 1196
Db 1530 CGAGTGCCTACTGATCGGTTTATCGGCGGCTGGACCAAGGAGGCGTGGACATCAT 1589
QY 1197 CTGGGCGCCCTGCCAAGATCTGCGCCACCCCAAGGTGCAGATGCCATCCTGGGTAC 1256
Db 1590 CGCGGAGCGATCCACTGGATC---GCGGGGAGGAGCTGAGCTSGTATGCTGGCAC 1646
QY 1257 CGCAGAGCGCCCTACGAGAGTGTGAACGCTATCGGCAAGTACAAAGGCGCGCC 1316
Db 1647 CGGCGGCGGCTGAGGAGATGTGCGGCGGTTCAGTCTGAGGACACAGCAAGGT 1706
QY 1317 CAAGGCGGTGTTCAAGTCTTCGCGCCCTGGCGGCATCTCACCGCGCGCGGCTT 1376
Db 1707 GCSCGCTGGTGGGTCTCTCGTGGCCCTGGCGCACCGCATCCGCGGCGCGGACAT 1766
QY 1377 CATGCTGGTCCCTCGGCTTCAGGCGCTGCGGCCCTGATCCAGCTGCACGCCATGCACTA 1436
Db 1767 CTGCTGATGCGCTGCGGCTTCAGCGCTGCGGCTGAACCACTCTACGCCATGGCGTA 1826
QY 1437 CGGTACCGTGGCGGTGAGCTTCCACCGCGCGGCTGTCGACACCGTCAAGAGGCGGT 1496
Db 1827 CGGACCGTGGCGGTGACCGCTTGGGGGGCTCCCGGACACAGCTGGCGCGCTTCA 1886
QY 1497 CACCGCTTCCATATGGGCGCTTGAACCCCGAAGCTGGAGAGGCTGACGCCGACGC 1556
Db 1887 CCGCTTCAACGACACCGGCTGCGGTGGAGTTCGACCGCGGAGCGAACCAGATGAT 1946
QY 1557 CTTGGCGGCGGCGGCTGCGAGCGAGTGTGCGGCGGCGGCTGACCCCGAGAT 1616
Db 1947 CGACGCGCTCTCGACTGCTCACCAGTACCGGAATACAGGAGAGCTGGCGCGCTG 2006
QY 1617 GGTGGCCAACTGATCAGCGAGGAGCTGCTGCTGCTCCAGCGCGGCGGAGTGGGAGG 1676
Db 2007 CAGGCGCGGCGGCTGCGGAGGAGCTGAGTGGGACCAAGCGCGGCTGCTGATGAGGA 2066
QY 1677 CTGCTGGAGGAGTGGTGTAC 1698
Db 2067 CTGCTGCTGAAGGCGAAGTAC 2088

RESULT 13
AAC86412
ID AAC86412 standard; cDNA; 2107 BP.
XX
AC AAC86412;
XX
DT 01-WAR-2001 (first entry)
XX Wheat starch synthase II coding sequence SEQ ID NO: 5.
DE Wheat; starch synthase II; SSIII; starch content; starch synthesis;
KW food product; adhesive; ss.
XX
OS Triticum aestivum.

XX WO200066745-A1.
PN 09-NOV-2000.
XX 28-APR-2000; 2000WO-AU00385.
XX 29-APR-1999; 99AU-0000052.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD.
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX
PI Morell M, Li Z, Rahman S, Appels R;
XX WPI; 2000-647602/62.
XX P-PSDB; AAB37568.
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
WST-II, useful in modifying plant starch content and/or composition -
XX Claim 1; Page 161-163; 21lpp; English.
XX The present invention relates to novel protein and coding sequences from
wheat. The proteins are wheat starch synthases, designated SSII and
SSIII. These can be used in the modification of plant starch content or
composition, and to screen plants to identify mutations which affect
starch content and composition. The starch can then be used in food
products, such as flour, and in films, coatings, adhesives, building
materials and packaging materials.
XX Sequence 2107 BP; 452 A; 566 C; 664 G; 425 T; 0 other;

Query Match 10.1%; Score 214; DB 21; Length 2107;
Best Local Similarity 55.8%; Pred. No. 8e-23;
Matches 478; Conservative 0; Mismatches 360; Indels 18; Gaps 3;
QY 858 CAACCTGCTGAAGGTTGGCAATTCGCGCGCGACAACTGCTGCTGCGCAACTA 917
Db 933 CAACCTACTTCGCGCGGCTGAAGATGCGGACAGGTTGCTGTGAGCGCCCGGTA 992
QY 918 CGCGACCGAGATCGTCCGATGCGCGCGGCTGTGAGCTGGACACCGCTATCC---G 974
Db 993 CTTGTGGAGCTGAAGACCGGTGGAGGCGCTTTCAGGACATCATACGGCAGAA 1052
QY 975 CGCCAAAGGGCATTTGAGGCGATTTGAACGCGATGAGAGTGGAAACCCAGAC 1034
Db 1053 CGACTGGAAGACCGCGGCACTGTCACGCGCATCGAACATGGAGTGAACCCGAGGT 1112
QY 1035 CGACAAGTCTCTGCTGCGCCCTTACGACCAAGACAGGCTCT-----ACGCGG 1082
Db 1113 GGACGCGGCTTCAAGTTCGAGCGCTTACCACTTCTCCCTGAGGACGCTGACTCCGG 1172
QY 1083 CAAAGCGCGCGCAAGGAGGCTTGCAGGCGGAGCTGGGCTGCTGTGACCCACCGC 1142
Db 1173 CAAAGCGGAGTGAAGGAGGCGCTGCGCGGAGCTGGGCTGCTGCTGCTGCTGCT 1232
QY 1143 CCGCTTCTGCTTCTGCTGCGCGCTTGGAGGAGAGAGGCTGTGAGATCATCTGCTGC 1202
Db 1233 GCGGCTGCTGCTTCTGCTGCGCGCTTGGAGGAGAGGCGTGGAGATCATCGGGA 1292
QY 1203 CGCGCTGCGCAAGATCTTGGCGACCCCAAGTTCAGATTCGCGCTCTGGTACCGGCAA 1262
Db 1293 CGCCATGCTGGATCGTG---AGCCAGGACGCTGAGCTGTGATGCTGGACCGGCG 1349
QY 1263 GCGCGCTACGAGAGCTGTGAACCCATCGCACCAAGTACAGGCGCGCGCAAGGG 1322
Db 1350 CCAGACCTTGAGAGCATGCTGACGACTTCGAGCGGAGACACACGACAAAGTGGCGCG 1409
QY 1323 CTTGTCTCAAGTTCCTGCGCGCTTGGCGCATGCTTACCGCGCGGCGGCTTCTGCT 1382
Db 1410 GTGGGTGGGTTCTCGTGGCGCTGGCGCACCGGATTCAGCGCGGCGGCGGCTCT 1469

QY 1383 GGTGCTCGCGCTTCGAGCCCTGGGCTGATCCAGCTCCAGCCATGCATCTACGGTAC 1442
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1470 CATGCCCTCCGGTTCGTGGCTGGGGCTGRACAGCTTACGCCATGSCCTACGGCAC 1529
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1443 CGTGCCTCGGTAGCTCCACCGCGCGCTGTGTCGACACCGCTCAAGAGGCGCTACCGG 1502
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1530 CGTCCCGCTCGTACAGCGCTGGCGGCTCAGGAGACCGGTGCCGCCCTTCGACCCCTT 1589
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1503 CTTCCATGGGCGCTTACGCCGACAGCTGGAGGCTGACGCCGACCGCCCTGGC 1562
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1590 CAACCATCCGGCTCGGGTGGAGCTTCGACCGCGCGGAGGCTCAAGGAGGCTCCAGGA 1649
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1563 CGCCACCGTGGCGCTGCCAGCGAGGTGTTGGCGGCGCGCTACCCCGAGATGTGGC 1622
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1650 GCTCGGCACTGCTCCGCACTACCGAGCTTCAAGGAGAGCTGGAGGSCCCTCCAGGA 1709
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1623 CAACCTCATCAGCAGAGCTGCTCGGTCCCAAGCGCGCGGCGCAGAGTGGAGGCGCTCT 1682
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1710 GCGCGCATGTCCAGGACTTACGCTGGGAGCACGCGCGCAAGCTCTACGAGGACGCTCT 1769
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1683 GGAGGAGGTGGTGTAC 1698
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1770 CGTCAAGGCCAAGTAC 1785
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||

RESULT 14

AAC86434
 ID AAC86434 standard; cDNA; 2107 BP.

XX AAC86434;

XX 01-MAR-2001 (first entry)

XX Wheat starch synthase clone wSSIIB cDNA.

XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
 KW food product; adhesive; ss.

XX Triticum aestivum.

XX WO20006745-A1.

XX 09-NOV-2000.

XX 28-APR-2000; 2000WO-AU00385.

XX 29-APR-1999; 99AU-0000052.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX (GOOD-) GOODMAN FIELDER LTD.

XX (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

XX Morell M, Li Z, Rahman S, Appels R;

XX WPI; 2000-647602/62.

XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 WST-II, useful in modifying plant starch content and/or composition -

XX Disclosure; Fig 2; 211pp; English.

XX The present invention relates to novel protein and coding sequences from
 CC wheat. The proteins are wheat starch synthases, designated SSII and
 CC SSIII. These can be used in the modification of plant starch content or
 CC composition, and to screen plants to identify mutations which affect
 CC starch content and composition. The starch can then be used in food
 CC products, such as flour, and in films, coatings, adhesives, building
 CC materials and packaging materials.

XX Sequence 2107 BP; 452 A; 566 C; 664 G; 425 T; 0 other;

XX Query Match 10.1%; Score 214; DB 21; Length 2107;
 XX Best Local Similarity 55.8%; Pred. No. 8e-23;

Matches 478; Conservative 0; Mismatches 360; Indels 18; Gaps 3;
 QY 858 CAACCTGCTCAAGGGTGGCAATTATCGCCGCGGACAAACCTGGTACTGTGTGCCCCAACTA 917
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 933 CAACCTACTTCGCGCGCGGCTGAAGATGGCGACCAAGTTGTGTGTGAGCCCGGGTA 992
 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 918 CGGACCGAGATCGTCCCGATGCGCGCGCGGTGTGGAGCTGGACACCGTATCC---G 974
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 993 CTTGTGGAGCTGAAGACGCTGGAGGCGGTGGGGGCTTCAGACATCATACGGCAGAA 1052
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 975 CGCAAGGGATTCAGGGCATTTGAACGCATGGACATTGAGGATGGAAACCCCAAGAC 1034
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1053 CGACTGAAGACCGCGGCATCGTCAACGGCATCGACAACATGAGGTGGAAACCCGAGGT 1112
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1035 CGACAAGTTCTGTCTGCGCCCTACGACCAAGACAGCTCT-----ACGCCG 1082
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1113 GGAGCCCACTCAAGTCGGACGGCTACCAACTTCTCCTGAGGACGCTGGACTCCG 1172
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1083 CAAGCGCGCGCCCAAGAGGCGCTTCGAGCGGAGCTGGGCTGCTGTGGACCCACCGC 1142
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1173 CAAGCGGACGTGCAAGAGGCGCTTCAGCGAGCTGGGCTGCGGTTCGCGCGGACGT 1232
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1143 CCCTCTGTCGCTTCATCGCGCCCTGGAGAGCAGAAGGTGTGGACATCATCTTGGC 1202
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1233 GCGCTGCTCGCTTCATCGCGCCCTGGACGGCGTGGACGAGGCGTGGAGATCATCGGGA 1292
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1203 CGCCCTGCCCAAGATCTTGCCACCCCAAGGTGCAGATGCCATCTTGGTACCGGCA 1262
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1293 CGCCATGCCCTGGATCGTG---AGCCAGGACGTGCAGTGTGTGATGTGGACCGGCG 1349
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1263 GGCGGCTACGAGAAGCTGTGAACGCCATCGGCACCAAGTACAAAGGCGCGCCCAAGGG 1322
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1350 CCACGACTGGAGAGCATGCTGCAGCACTTCGAGCGGAGACCAACGACAAAGGTGCGCGG 1409
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1323 CGTGTCAAGTCTTCGCGCCCTGGCGCACATGTCACCGCGCGCGCGGCTTCTATGCT 1382
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1410 GTGGTGGGTTCCTCGTCCCTGGCGCACCGCATACGCGGGGGGCGGACGCTCT 1469
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1383 GGTGCTCGCGCTTCGAGCCCTGCGGCTGATCCAGTGCACGCCATGCACTACGGTAC 1442
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1470 CATGCCCTCCGGTTCGTGCGGTGCGGCTGAACGAGCTTACGCCATGGCTACGGCAC 1529
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1443 GGTGCGGCTGTAGCTCCACCGCGGCTGTGCACACCTCAAGGAGGCGCTCACCGG 1502
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1530 GGTCCCGTGTGTCACGCGCTTCAGGACACCGCTGCGCGCTTCGCGGCTTCGACCCCTT 1589
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1503 CTTGCACATGGCGCGCTGAACCCCGCAAGCTGCACAGGCTGACGCCGACGCCCTGGC 1562
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1590 CAACCACTCCGGGCTCGGTGGACCTTCGACCGCGCGGAGGCGCACAGCTGATCGAGGC 1649
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1563 CGCCACCGTGGCGCTGCCAGCGAGGTGTTTGGGGCGGCGGCTACCCCGAGATGGTGGC 1622
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1650 GCTGGGCACTGCTTCGCACTACCGAGACTTCAAGGAGAGCTGGAGGCGCTCCAGGA 1709
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1623 CAACCTCATCAGCAGGACCTGCTGTGTCGAAGCCCGCCAGAGTGGAGGCGCTGCT 1682
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1710 GCGGCGCATGTGCGAGGACTTACGCTGGGAGCACCCCGCCCAAGCTCTACGAGGACGCTCT 1769
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 QY 1683 GGAGGAGGTGGTGTAC 1698
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
 Db 1770 GGTCAAGGCCAAGTAC 1785
 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||

RESULT 15

AAV66834
 ID AAV66834 standard; DNA; 2380 BP.

XX AAV66834;

XX 05-JAN-1999 (first entry)

XX Zea mays soluble starch synthase gene SSS56.

Db 2141 ARTAYCARTGGTTRGONAAATTRYTNGCNACNMGNGNWSNTGYMGNMGNACNTGGA 2200
QY 1752 CGAGAAAGATCCCCGGCGACCTGCCCCCGGTGTCCTACGCCCCCAACACCCCTGAAGCC 1808
Db 2201 CNYTNTTYMGNMGNTTNTTWSNYTNGCNCNTYTNATGMGNGCNSNCAYTTNMGNM 2257

Search completed: June 3, 2003, 10:07:07
Job time : 351.686 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:33:29 ; Search time 2072.3 Seconds
(without alignments)
16599.542 Million cell updates/sec

Title: US-09-980-771a-2

Perfect score: 2124

Sequence: 1 atgctgttgctctaccag.....cgccaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508.6	23.9	786	12	BF864001 963048D01
2	503.6	23.7	511	10	AV394077 AV394077
3	476.6	22.4	502	14	BQ814687 BQ814687
4	469.8	22.1	473	10	AV640741 AV640741
5	466.8	22.0	506	10	AV643281 AV643281
6	458	21.6	475	10	AW757933 AW757933

7	457.8	21.6	461	10	AV629488
8	456.8	21.5	2147	11	AY109531
9	456	21.5	666	13	BM003295
10	444.4	20.9	458	10	AV622787
11	443.8	20.9	560	10	AV628313
12	432.8	20.4	436	10	AV629117
13	422.8	19.9	552	10	BE024926
14	421.2	19.8	517	10	AV641583
15	415.2	19.5	532	10	AV641989
16	414.8	19.5	418	10	AV644765
17	406.8	19.2	513	10	AV631004
18	406.8	19.2	552	10	AV640546
19	405.8	19.1	524	10	AV641724
20	405.2	19.1	500	10	AV642869
21	402.8	19.0	510	10	AV395307
22	394.8	18.6	540	10	AV642722
23	391.4	18.4	393	10	AV631115
24	382.8	18.0	486	10	AV642834
25	382.8	18.0	504	10	AV629198
26	381.8	18.0	502	10	AV644517
27	379.8	17.9	495	10	AV642565
28	371.8	17.5	493	10	AV631838
29	370.8	17.5	492	10	AV629891
30	369.8	17.4	484	10	AV629266
31	369.8	17.4	489	10	AV629190
32	364.8	17.2	487	10	AV643763
33	364.6	17.2	385	10	AV393664
34	359.8	16.9	490	10	AV628140
35	358.8	16.9	476	10	AV644278
36	356.8	16.8	478	10	AW758015
37	356.8	16.8	502	10	AV631098
38	354.8	16.7	474	10	AV397080
39	353.8	16.7	468	10	AV644593
40	352.8	16.6	474	10	AV630110
41	352.6	16.6	367	13	BI727902
42	351.8	16.6	487	10	AV629387
43	351.2	16.5	799	12	BF863935
44	330.2	15.5	453	10	AV643760
45	326.8	15.4	448	10	AV644481

ALIGNMENTS

RESULT 1
BF864001 786 bp mRNA linear EST 19-JAN-2001
LOCUS 963048D01.yl C. reinhardtii CC-1690, Stress condition I, normalized
DEFINITION , Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BF864001
VERSION BF864001.1 GI:12254145
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
AUTHORS Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C., Lefebvre, P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; Project phase 3
JOURNAL Unpublished (2000)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
FEATURES
Location/Qualifiers
source 1..786
/organism="Chlamydomonas reinhardtii"

AV629488 AV629488
AY109531 Zea mays
BM003295 1031109F0
AV622787 AV622787
AV628313 AV628313
AV629117 AV629117
BE024926 894006H08
AV641583 AV641583
AV641989 AV641989
AV644765 AV644765
AV631004 AV631004
AV640546 AV640546
AV641724 AV641724
AV642869 AV642869
AV395307 AV395307
AV642722 AV642722
AV631115 AV631115
AV642834 AV642834
AV629198 AV629198
AV644517 AV644517
AV642565 AV642565
AV631838 AV631838
AV629891 AV629891
AV629266 AV629266
AV629190 AV629190
AV643763 AV643763
AV393664 AV393664
AV628140 AV628140
AV644278 AV644278
AW758015 874005B12
AV631098 AV631098
AV397080 AV397080
AV644593 AV644593
AV630110 AV630110
BI727902 1031095E0
BF863935 963047G01
AV643760 AV643760
AV644481 AV644481

```
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, lambda Zap II"
/Note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
Zap clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT 157 a 282 c 244 g 102 t 1 others
ORIGIN

Query Match 23.9%; Score 508.6; DB 12; Length 786;
Best Local Similarity 95.3%; Pred. No. 5e-78;
Matches 546; Conservative 0; Mismatches 24; Indels 3; Gaps 2;

QY 1554 CGCCCTGGCGCCACCGCTGGCGGTGCCAGCGAGGTGTTTGGCGGGCGCGCTTACCCCGA 1613
Db 95 CCCCCGGCTGCAGGAATTCGGCACGAGGCGAGGTGTTTGGCGGGCGCGCTTACCCCGA 154

QY 1614 GATGTGGGCACTGATCAGCAGCAGACCTGCTCTGGTCCAGCCCGCCGAGAGTGGGA 1673
Db 155 GATGTGGGCACTGATCAGCAGCAGACCTGCTCTGGTCCAGCCCGCCGAGAGTGGGA 214

QY 1674 GGCCCTGCTGGAGAGGTGTGTACGGCAAGGCGCGGTGGCCACCGCCCAAGAGGAGGA 1733
Db 215 GGGCCCTGCTGGAGAGGTGTGTACGGCAAGGCGCGGTGGCCACCGCCCAAGAGGAGGA 274

QY 1734 GATCAAGGTGCCCGTTGCCGAGAGATGCCCGGCGACCTGCCCGCGGTGTCTTACGCCCC 1793
Db 275 GATCAAGGTGCCCGTTGCCGAGAGATGCCCGGCGACCTGCCCGCGGTGTCTTACGCCCC 334

QY 1794 CAACACCTGAAGCCGTGTCGGCTCCGTTGGAGGGCAAGCGCGCGCCGCGCCCAAGGT 1853
Db 335 CAACACCTGAAGCCGTGTCGGCTCCGTTGGAGGGCAAGCGGTGCGCGCGCCGCGCCCAAGGT 394

QY 1854 CGCACACACCGCCCGCCATGGCGGTGGCGGACACACCCCTCGGCGCCCGCTCGCC 1913
Db 395 CGCACACACCGCCCGCCATGGCGGTGGCGGACACACCCCTCGGCGCCCGCTCGCC 454

QY 1914 CGCCGGCGCCACCCCGCAAGGTGACACCTACAGCCCGCGCCCTGCGCCGACCGCCAGCC 1973
Db 455 CGCCGGCGCCACCCCGCAAGGTGACACCTACAGCCCGCGCCCTGCGCCGACCGCCAGCC 514

QY 1974 CAAGACCGCTGGCTCAAGTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGG 2033
Db 515 CAAGACCGCTGGCTCAAGTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGG 574

QY 2034 CGCTGCTCCCAAGCGCAACGGCAACGGTGGCTCGGCTCCCAAGACCTCGGCTGCCAAGCC 2093
Db 575 CGCTGCTCCCAAGCGCAACGGCAACGGTGGCTCGGCTCCCAAGACCTCGGCTGCCAAGCC 633

QY 2094 CCTGGT--CTCGCGCGCCACCGCAAGTCCGCC 2124
Db 634 CCTGGTCTTCGGGGGCCACCGGAAGTTCGCC 666

RESULT 2
AV394077 511 bp mRNA linear EST 23-APR-2002
LOCUS AV394077 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION cDNA clone CL18c08_r 5', mRNA sequence.
ACCESSION AV394077
VERSION AV394077.1 GI:6548293
KEYWORDS EST.
```

Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 511)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. 1. Generation of 3433
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)
20152988
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..511
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone_lib="CL18c08_r"
/clone_lib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 80 a 170 c 159 g 101 t 1 others
ORIGIN

Query Match 23.7%; Score 503.6; DB 10; Length 511;
Best Local Similarity 99.0%; Pred. No. 3.4e-77;
Matches 506; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 64 TTCGGTGTCAAGAGACCGCGAACCGCTGCTGCGTGAGCTTGCTGCGTCGCGACGC 123
Db 1 TTCGGTGTCAAGAGACCGCGAACCGCTGCTGCGTGAGCTTGCTGCGTCGCGACGC 60

QY 124 AAGTCCACCTCGCGCTCGGCTGTTACTGTGTCACCTGCTGCCACTGCGCGCTGGACATC 183
Db 61 AAGTCCACCTCGCGCTCGGCTGTTACTGTGTCACCTGCTGCCACTGCGCGCTGGACATC 120

QY 184 GTGATGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCGGCGCTGGCGATGTGACT 243
Db 121 GTGATGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCGGCTGGCGATGTGACT 180

QY 244 GGTGGCTGCGCTATTGAGCTGTGTCAGCGGGGCCACCGCGTCATGACCATTTGCCCTCGC 303
Db 181 GGTGGCTGCGCTATTGAGCTGTGTCAGCGGGGCCACCGCGTCATGACCATTTGCCCTCGC 240

QY 304 TACGACAGTACGCTGAGCTGGACACCTCGGTGTCGTCGACATCATGGCGAGAG 363
Db 241 TACGACAGTACGCTGAGCTGGACACCTCGGTGTCGTCGACATCATGGCGAGAG 300

QY 364 GTCCGCTACTTCCACTCCATCAAGAAGGCGTGCACCGCGTGTGGATTGACACCCCTGG 423
Db 301 GTCCGCTACTTCCACTCCATCAAGAAGGCGTGCACCGCGTGTGGATTGACACCCCTGG 360

QY 424 TTCTGGCCAAAGGCTTGGGGCAAGACCGGTTCGAAGCTGTACGGCCCGCGCTCGCGGCT 483
Db 361 TTCTGGCCAAAGGCTTGGGGCAAGACCGGTTCGAAGCTGTACGGCCCGCGCTCGCGGCT 420

QY 484 GACTACCTGGACACACCAAGCGCTTCCCTGCTTCTCAAGGCGCGCTATTGAGCTGCC 543
Db 421 GACTACCTGGACACACCAAGCGCTTCCCTGCTTCTCAAGGCGCGCTATTGAGCTGCC 480

QY 544 CGCGTGTGCCCTTCGGCCCGCGGAGGACT 574
Db 481 CGCGTGTGCCCTTCGGCCCGCGGAGGACT 511

RESULT 3
B0814687 502 bp mRNA linear EST 01-AUG-2002
LOCUS B0814687

[illegible]

```

Db 241 TGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGCTGTACGGGCCCGCTCCGGCGCTGACT 300
QY 488 ACCTGGACAACACAAAGCGCTTGGCCCTGTTCTGCTCAAGCGCCGCTATTAGGCTGCCGGG 547
Db 301 ACCTGGACAACACAAAGCGCTTGGCCCTGTTCTGCTCAAGCGCCGCTATTAGGCTGCCGGG 360
QY 548 TGCTGCGCTTTCGGCCCGCGGAGGAGTGGCTTCTGCTGCGCCACGACTGGCACTCCGGCC 607
Db 361 TGCTGCGCTTTCGGCCCGCGGAGGAGTGGCTTCTGCTGCGCCACGACTGGCACTCCGGCC 420
QY 608 TGGTGGCGCTTCTGCTGAAGGACGAGTACCAAGCCCAAGGCGCGAGTTCACCAAG 660
Db 421 TGGTGGCGCTTCTGCTGAAGGACGAGTACCAAGCCCAAGGCGCGAGTTCACCAAG 473

RESULT 5
AV643281 506 bp mRNA linear EST 15-DEC-2000
LOCUS AV643281 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION cDNA clone HCL066h09_r 5', mRNA sequence.
ACCESSION AV643281
VERSION AV643281.1 GI:10786609
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Grossman, A., Davies, J., Inoue, Y., Fukuzawa, H., Ohya, K.,
AUTHORS Asamizu, E., Miura, K., and Tabata, S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
JOURNAL adapted cells of Chlamydomonas reinhardtii
MEDLINE DNA Res. 7 (5), 305-307 (2000)
COMMENT 20539644
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/eu/plant/.

FEATURES
source
1. 506
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="HCL066h09_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT 77 a 170 c 160 g 99 t
ORIGIN
Query Match 22.0%; Score 466.8; DB 10; Length 506;
Best Local Similarity 99.6%; Pred. No. 7.5e-71;
Matches 468; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCGCTGCTGCTACAGCGCCCGCAGCAGCGCGTCTCTATCGTCATCAATGCCGCG 60
Db 37 ATGCGCTGCTGCTACAGCGCCCGCAGCAGCGCGTCTCTATCGTCATCAATGCCGCG 96
QY 61 TCCTGCTGCTGCTCAAGAACCGCGAACAGCTGCTGCTGAGCTTGCTGCTGCTCGCA 120
Db 97 TCGTTGCTGCTCAAGAACCGCGAACAGCTGCTGCTGAGCTTGCTGCTGCTCGCA 156
QY 121 CGCAAGTCCAGCTCGCGCTGGCTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 157 CGCAAGTCCAGCTCGCGCTGGCTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
QY 181 ATCGTGATGTTGCTGCTGAGCTGCGCCCTTGGTCCCAAGACGGCGGCTCGGCGATGTG 240
Db 217 ATCGTGATGTTGCTGCTGAGCTGCGCCCTTGGTCCCAAGACGGCGGCTCGGCGATGTG 276

```

```

QY 241 ACTGTGGCGCTGCCTATTGAGCTGTGTCAGCGCGGCACCGCGCTCATGACATTGCCCT 300
Db 277 ACTGTGGCGCTGCCTATTGAGCTGTGTCAGCGCGGCACCGCGCTCATGACATTGCCCT 336
QY 301 CGCTAGCACAGTACGCTGAGCGCTGGGACACCTCGGTGCTGTCGACATCATGGCGAG 360
Db 337 CGCTAGCACAGTACGCTGAGCGCTGGGACACCTCGGTGCTGTCGACATCATGGCGAG 396
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACACCCC 420
Db 397 AAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATTGACACCCC 456
QY 421 TGGTTCTTGGCCAAAGTCTGGGCAAGACCGGCTCCAAAGCTGTAGCGGCC 470
Db 457 TGGTTCTTGGCCAAAGTCTGGGCAAGACCGGCTCCAAAGCTGTAGCGGCC 506

RESULT 6
AW757933 475 bp mRNA linear EST 03-MAY-2000
LOCUS 874004E12.y1 C. reinhardtii CC-1690, Lambda Zap II Chlamydomonas
DEFINITION reinhardtii cDNA, mRNA sequence.
ACCESSION AW757933
VERSION AW757933.1 GI:7687285
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,
AUTHORS McDermott, J.P., Silflow, C., Stern, B. and Surzycki, R.,
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
JOURNAL Unicellular System for Analyzing Gene Function and Regulation in
COMMENT Vascular Plants: project phase 2
Unpublished (2000)
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlany@duke.edu.

FEATURES
source
1. 475
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage."
BASE COUNT 74 a 160 c 143 g 98 t
ORIGIN
Query Match 21.6%; Score 458; DB 10; Length 475;
Best Local Similarity 97.9%; Pred. No. 2.4e-69;
Matches 464; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 77 AGACGGCAACACAGCTGCTGCTGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 136
Db 2 AGACGGCAACACAGCTGCTGCTGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 61
QY 137 GCTCGCTGTTTACTGCTGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 196

```



```

Db 62 GCTGGCTGTACTGTCGCCACTGGTGCACCTTGTCTGCTGGACATCTGTGAIGTTGCTG 121
QY 197 CTGAGGTGCGCCCTTGTTCAGAGAGGGCGCCCTGGGGATGTGACTGGTGGCCCTGCGCTA 256
Db 122 CTGAGGTGCGCCCTTGTTCAGAGAGGGCGCCCTGGGGATGTGACTGGTGGCCCTGCGCTA 181
QY 257 TTGAGCTGTGTCAGAGCGGGCGCCACCGGTCATGACCATTTGCCCTGCTACGACCACTAG 316
Db 182 TTGAGCTGTGTCAGAGCGGGCGCCACCGGTCATGACCATTTGCCCTGCTACGACCACTAG 241
QY 317 CTGAGCGCTGGGACACCTCGGTGTGCTGGACATCATGGGCGAGAGGTTCGCTACTTCC 376
Db 242 CTGAGCGCTGGGACACCTCGGTGTGCTGGACATCATGGGCGAGAGGTTCGCTACTTCC 301
QY 377 ACTCCATCAAGAGGGCGTCACCGGTCATGACCATTTGCCCTGCTACGACCACTAG 436
Db 302 ACTCCATCAAGAGGGCGTCACCGGTCATGACCATTTGCCCTGCTACGACCACTAG 361
QY 437 TCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCCGCTCGGGCTGACTACCTGGACA 496
Db 362 TCTGGGCTAGACCGGCTCCAAAGCTGTACGCGCCCGCTCGGGCTGACTACCTGGACA 421
QY 497 ACCACAGCGCTTGGCCCTCTTCTGCAAGCGCGCTATTGAAGCTGCGCGGTGC 550
Db 422 ACCACAGCGCTTGGCCCTCTTCTGCAAGCGCGCTATTGAAGCTGCGCGGTGC 475

RESULT 7
LOCUS AV629488 461 bp mRNA linear EST 15-DEC-2000
DEFINITION AV629488 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
reinhartdii cDNA clone LCL059f09_r 5', mRNA sequence.
ACCESSION AV629488
VERSION AV629488.1 GI:10792122
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 461)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..461
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL059f09_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
/note="Vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%."
BASE COUNT 73 a 156 c 143 g 89 t
ORIGIN
Query Match 21.6%; Score 457.8; DB 10; Length 461;
Best Local Similarity 99.6%; Pred. No. 2.6e-69;
Matches 459; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 179 ACATCGTATGTTGCTGCTAGGTGCGCCCTTGTCTCAAGACGGCGCTGGCGGATG 238
|||||

```

```

Db 1 ACATCGTATGTTGCTGCTAGGTGCGCCCTTGTCTCAAGACGGCGCTGGCGGATG 60
QY 239 TGACTGTGGCTGCTGCTATTGAGCTGGTCAAGCGGGCGCCACCGGCTCATGACCATGGCC 298
Db 61 TGACTGTGGCTGCTGCTATTGAGCTGGTCAAGCGGGCGCCACCGGCTCATGACCATGGCC 120
QY 299 CTCGTACGACCACTGCTGCTGACGCTGGGACACCTGCTGCTGGTGTGACATCATGGCG 358
Db 121 CTCGTACGACCACTGCTGCTGACGCTGGGACACCTGCTGCTGGTGTGACATCATGGCG 180
QY 359 AGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATTGACCAACC 418
Db 181 AGAAGTCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATTGACCAACC 240
QY 419 CCTGTTCTCTGCGCAAGGTCTGGGGCAAGACCGCTCCCAAGCTGTACGCGCCCGCTCCG 478
Db 241 CCTGTTCTCTGCGCAAGGTCTGGGGCAAGACCGCTCCCAAGCTGTACGCGCCCGCTCCG 300
QY 479 GCGTGTACTGCTGACCAACCAAGCGCTTCGCGCTTCTGCAAGCGCTATTGAGG 538
Db 301 GCGTGTACTGCTGACCAACCAAGCGCTTCGCGCTTCTGCAAGCGCTATTGAGG 360
QY 539 CTGCCCCGGTGTGCGCTTCGCGCCCGCGGAGGACTGCGCTTTCGTGGCCCAAGCACTGGC 598
Db 361 CTGCCCCGGTGTGCGCTTCGCGCCCGCGGAGGACTGCGCTTTCGTGGCCCAAGCACTGGC 420
QY 599 ACTCCGCGCTGCTGCGCGCTCTGCTGAAGGACGAGTACCAG 639
Db 421 ACTCCGCGCTGCTGCGCGCTCTGCTGAAGGACGAGTACCAG 461

RESULT 8
LOCUS AY109531 2147 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays CUL198_1 mRNA sequence.
ACCESSION AY109531
VERSION AY109531.1 GI:21213285
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2147)
AUTHORS Rainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2147)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
source
1..2147
/organism="Zea mays"
/db_xref="MaizeDB:630100"
/db_xref="taxon:4577"
/clone="CUL198_1"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project."
BASE COUNT 400 a 582 c 604 g 365 t 196 others
ORIGIN
Query Match 21.5%; Score 456.8; DB 11; Length 2147;
Best Local Similarity 58.5%; Pred. No. 4.7e-69;

```

	Matches	917; Conservative	0; Mismatches	555; Indels	96; Gaps	6
QY	161	GTGCCATTCGGCGCTGGACATCGTGAATGGTTGCTGTAGGTGCCCGCTTGGTCCAAAGA	220			
Db	215	GCGCCAGCGCGGATGAACGTCTGCTCGCGCGGAGATGCGCGGTGGAGCAAGA	274			
QY	221	CGGCGGCGCTGGGGATGTGACTGGTGGCCCTGCTATTGACTGGTCAAGCGCGGCCACC	280			
Db	275	CCGCGGCGCTCGGCGACGNN	334			
QY	281	GCCTCATGACATTCGCCCTCGCTACGACCAAGTACGTACGCTTGGGACACCTTCGGTGG	340			
Db	335	GTGCTATGGTCTCTCTCCCGCTACGACCAAGTACAGGACGCTTGGGACACCAAGCTG	394			
QY	341	TCGTGGACATCA-----TGGCGGAGAAGTCCGCTACTTCACATCCATCAAGA	388			
Db	395	TGTCGGAGATCAAGATGGGAGACGGTACGAGACGGTCAAGTCTTCTTCCACTGCTACAAGC	454			
QY	389	AGGCGTGCACCGGTGTGGATTACCAACCCCTGGTCTCTGCGCAAGGTCTGGGCAAGA	448			
Db	455	CGGAGTGGACCGGTGTCTGTACCAACCCACATGTTCTTGAGAGGNNNNNNNNNNNNNN	514			
QY	449	CCGCTCTCAAGCTTACGGCCCGCGCTCCGCGCTGACTACTCTGTCACCAACCAAGCGCT	508			
Db	515	NN	574			
QY	509	TCGCCCTGTCTGCAAGGCCGCTATTGAGGTGCTCCCGGTGCTGCCCTTCGGC-----	561			
Db	575	TCAGCCTGCTATGCGGACGACACTTGAAGTTCGAAGTATCCAGGACCTTACGCTTCAACA	634			
QY	562	-----CCCGCGGAGGACTGCGTCTTCTGTCGCAACGACTGCGACTCCGCTCG	604			
Db	635	CATACTTCTCGGACCATACGGGAGGACGTCGTCTGCTGCAACGACTGGCACACCG	694			
QY	605	CCCTGGTCCCGTCTGCTGAAGGACGAGTACCACCCCAAGGGCCAGTTTACCAGGCCA	664			
Db	695	GCCCTCTCTGCTGTACTCTAAGAGCACTACACGCTCCACCGCATCTACAGGGAGCGAA	754			
QY	665	AGTCGGTCTGGCTATCCACAACATCGCTTCCAGGGCGCGATGTGGGAGGAGGCTTCCA	724			
Db	755	AGACCGTTTCTGATCCACAACATCTCTTACCAGGGCGGTTGCCCTTCTCCGACTACC	814			
QY	725	AGGACACGAAGTGTCCCAGGCGCGCTTTGACAAGCTGGCCTTCTGGACGGCTATGCCA	784			
Db	815	CGAGCTGAACCTCCCGGAGAGATTCAAGTCTGCTCTCGATTTCATCGACGGCTACGAGA	874			
QY	785	AGGTTACTACTGAGGCCACCCCATGCGAGGAGGACGAGAAGCCCGCTGACGGGAAAGA	844			
Db	875	AG-----CCGCTGGAAGG-----CCGCTGGAAGG-----	887			
QY	845	CCTACAAGAATCAACTGGCTGAAGGTTGGCATTCGCCCGCCGACAAAGCTGGTGA	904			
Db	888	---CCGGAAGATCAACTGGATGAAGCGCGGATCTCTGAGCGCCGACAGGCTCTCACCG	943			
QY	905	TGTCGCCCACTACGCGCCGAGATCGCTCCGATGCCCGCGCGCTGTGGAGCTGGACA	964			
Db	944	TCAGCCCTTACTACCGCGAGGAGCTATCTCCGGCATCGCCAGGGGCTGGAGCTTCACA	1003			
QY	965	CGCTCATCCGCCCAAGGSCATTGAGGCAATGTGAACGGCATGGAATTTGAGAGTGGGA	1024			
Db	1004	ACATATGGCTTACCGGSCATCACCGGCATGCTCAACGGCATGGAGCTCAGCGAGTGGG	1063			
QY	1025	ACCCCAAGACCGCAAGTTCTGCTGCGCCCTACACACAGAACAGCGTCTAGCCGGCA	1084			
Db	1064	ACCCCAAGGAGCAAGTATCGCGGTGAAGTACCAAGTGTGACGCGCGCTGGAGGCCA	1123			
QY	1085	AGGCGCGCCCAAGGAGGCCCTGCAGCGCGAGTGGGCTGCTGTGGACCCCAACCGCC	1144			
Db	1124	AGGCGTGAACAAGGAGGCGCTGCAGCGGAGTGGCGGCTCCCGGTGGACCGGAACATCC	1183			
QY	1145	CCCTGTTCGCTTCATTCGCGCGCCTTGGAGAGCAGAGGTTGTGGACATCATCTTGGCC	1204			
Db	1184	CGCTGTGGCGTTTCATTCGCGAGCTGGAGAGCAGAGGCGCCCGCAGTCTATGCGCGCC	1243			

QY	1205	CCCTGCCAAGATCTCTGGCCACCCCAAG---GTCCAGATCGCATCTCTGGGTACGGCA	1266
Db	1244	CCATCCGCGAGCTCAATGGAGATGGTGGAGAGCTGCATCGTTCTCTGGGCAAGGGCA	1303
QY	1262	AGCGCCCTACGAGAAGCTGGTGAAGCCATCGGCAACAACTACAAAGGCGCGCCCAAGG	1321
Db	1304	AGAGAAGTTTCGAGCGCATGCTCATGAGCGCGAGGAGAAGTTCAGGCAAGTGGCGG	1363
QY	1322	GCCTGGTCAAGTTCTCGCGCCCTTGGCGCACATGCTCACGCGCGCGCGCACTTCATGC	1381
Db	1364	CCGTGGTCAAGTTCAACGCGCGCTGGCGCACCATCATGCGCGCGCGCGCACTGCTCG	1423
QY	1382	TGCTGCGCCCTCGCGCTTCGAGCCCTCGGGCGCTGATCCAGCTGCAGCCCATGCACGGTA	1441
Db	1424	CCGTACCAAGCGCTTCGAGCCCTCGGGCGCTCATCCAGCTGCAGGGATGGGATACAGAA	1483
QY	1442	CCGTGCGCGTGTAGCTCTCACCGGGCGCTTGGTGCAGCACCGTCAAGGAGGCGCTCACCG	1501
Db	1484	CGCCCTCGCGCTCGGGTTCACCGGTGAGCTGCTGCACACCATCATCGAAGCAAGACCG	1543
QY	1502	GCCTTCCACATGGCGCCCTGA-----ACCCGACAAAGTGGACGAGCTGACGCCGAGG	1555
Db	1544	GGTTCACATGGCGCGCTCAGCGTGCAGTGAACGCTGTGGAGCGCGCGACGCTCAAGA	1603
QY	1556	CCCTGGCGCGCAGCTGGCGCGCTGCAGGAGGTGTTTCGGGCGCGCTACCCCGAGA	1615
Db	1604	AGTGGCGCACCACTTGCAGCGCGCATCAAGGTGTGCGGACGCCGCGCTACGAGGAGA	1663
QY	1616	TGGTGGCAACTGCATCAGCAGGAGCTGTCTCTGGTCCAAAGCGCCCGCAGAGTGGGAGG	1675
Db	1664	TGGTGGCAACTGCATCAGCAGGAGTCTCTCTGGAGGCGCTTCCCAAGAACTGGGAGA	1723
QY	1676	GCCTGCTG 1683	
Db	1724	ACGTGCTG 1731	
RESULT 9			
BM003295			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

```

1. .666
source
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 2lgr"
/db_xref="taxon:3055"
/clone.lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda zap II"
/note=vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John

```

Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH₄⁺ - containing) and shifted to TAP - NO₃⁻ (24hrs); H2 production conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O₂ (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806." 77 t

BASE COUNT 121 a 244 c 224 g 77 t
ORIGIN
Query Match 21.5%; Score 456; DB 13; Length 666;
Best Local Similarity 100.0%; Pred. No. 5.6e-69;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1669 TGGAGGCGCTGCTGAGAGGTGGTGTACGGCAAGGGGGCGCTGGCCACCGCAAGAAG 1728
Db 1 TGGAGGCGCTGCTGAGAGGTGGTGTACGGCAAGGGGGCGCTGGCCACCGCAAGAAG 60
QY 1729 GAGGAGATCAAGTGGCGCTTGCAGAGATCCCGGGGACCTGCCCGCGTGTCTTAC 1788
Db 61 GAGGAGATCAAGTGGCGCTTGCAGAGATCCCGGGGACCTGCCCGCGTGTCTTAC 120
QY 1789 GCGCCCAACACCTGAAGCCCGTGTCCGCTCCGTGGAGGCAACGGCGCGCGGCC 1848
Db 121 GCGCCCAACACCTGAAGCCCGTGTCCGCTCCGTGGAGGCAACGGCGCGCGGCC 180
QY 1849 AAGGTGGGACCAACCGCCCGCCATGGCGGTGGCGCGGACACCCCTCGGGGCC 1908
Db 181 AAGGTGGGACCAACCGCCCGCCATGGCGGTGGCGCGGACACCCCTCGGGGCC 240
QY 1909 TCGCCCGCGCGCCACCGCCCAAGGTGACACCTACAAGCCGCGCTGCCCGCCACCGCC 1968
Db 241 TCGCCCGCGCGCCACCGCCCAAGGTGACACCTACAAGCCGCGCTGCCCGCCACCGCC 300
QY 1969 AAGCCCAAGCGCTGGCCCTCAAGCTGGCGGTGAGGCTCCACACCTCGACCTCGGAG 2028
Db 301 AAGCCCAAGCGCTGGCCCTCAAGCTGGCGGTGAGGCTCCACACCTCGACCTCGGAG 360
QY 2029 AAGCGCGCTGCTCCCAAGCGACCGACGCTGCTGGCGCTCCAGACCTCGGCTGCC 2088
Db 361 AAGCGCGCTGCTCCCAAGCGACCGACGCTGCTGGCGCTCCAGACCTCGGCTGCC 420
QY 2089 AAGCCCGCTGGCTCGCGCGCCACCGCCCAAGTCCGCC 2124
Db 421 AAGCCCGCTGGCTCGCGCGCCACCGCCCAAGTCCGCC 456

RESULT 10
AV622787 458 bp mRNA linear EST 15-DEC-2000
LOCUS AV622787 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
DEFINITION reinhardtii cDNA clone LC054e04_r 5', mRNA sequence.
ACCESSION AV622787
VERSION AV622787.1 GI:10771964
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
AUTHORS Asamizu, E., Miura, K., Kuchino, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
JOURNAL adapted cells of Chlamydomonas reinhardtii
MEDLINE DNA Res. 7 (5), 305-307 (2000)
20539644

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research.
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1. 458
/organism="Chlamydomonas reinhardtii"
/strain="Cg"
/db_xref="taxon:3055"
/clone="LC054e04_r"
/notes="Vector: pBluescriptII SK-; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%." 74 t

BASE COUNT 91 a 147 c 146 g 74 t
ORIGIN
Query Match 20.9%; Score 444.4; DB 10; Length 458;
Best Local Similarity 99.6%; Pred. No. 5.3e-67;
Matches 456; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 536 AGGTGCGCGCTGCTGCCCTTGGCGCCCGGCGAGGACTCGTCTTCGTGGCCCAACGACT 595
Db 1 AGGTGCGCGCTGCTGCCCTTGGCGCCCGGCGAGGACTCGTCTTCGTGGCCCAACGACT 60
QY 596 GGCACCTCGCGCTGCTGCCCTGCTGAAGACAGTACCAGCCCAAGGGCCAGTTCA 655
Db 61 GGCACCTCGCGCTGCTGCCCTGCTGAAGACAGTACCAGCCCAAGGGCCAGTTCA 120
QY 656 CCAAGGCCAAGTGGTGGCTATC-CACAACTCCCTTCCAGGCGCCCATGTGGAG 714
Db 121 CCAAGGCCAAGTGGTGGCTATCGCAACATCGCTTCCAGGCGCCCATGTGGAG 180
QY 715 GAGGCTTTCAGGACAGAGCTGCCCGCCGCTTTGACAAGCTGGCTTCTCGGAC 774
Db 181 GAGGCTTTCAGGACATGAAGCTGCCCGCCGCTTTGACAAGCTGGCTTCTCGGAC 240
QY 775 GGTATGCCAAGGTTTACACTGAGGCGCCACCCCATGGAGGAGACGAGAGCCCCGCTG 834
Db 241 GGTATGCCAAGGTTTACACTGAGGCGCCACCCCATGGAGGAGACGAGAGCCCCGCTG 300
QY 835 ACGGGAAGACCTTACAAAGATCAACTGGCTGAAGGTGGCATTTATCGCGCCGCAAG 894
Db 301 ACGGGAAGACCTTACAAAGATCAACTGGCTGAAGGTGGCATTTATCGCGCCGCAAG 360
QY 895 CTGGTCACTGTGCGCCCAACTACGCGACCGAGATCGCTGCCGATCCCGCGCGGTGTG 954
Db 361 CTGGTCACTGTGCGCCCAACTACGCGACCGAGATCGCTGCCGATCCCGCGCGGTGTG 420
QY 955 GAGCTGCACACCGTCAATCCGCGCCCAAGGGCATTGAGGG 992
Db 421 GAGCTGCACACCGTCAATCCGCGCCCAAGGGCATTGAGGG 458

RESULT 11
AV628313 560 bp mRNA linear EST 15-DEC-2000
LOCUS AV628313 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
DEFINITION reinhardtii cDNA clone LCL039h11_r 5', mRNA sequence.
ACCESSION AV628313
VERSION AV628313.1 GI:10790947
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
AUTHORS Asamizu, E., Miura, K., Kuchino, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2

JOURNAL adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

Location/Qualifiers

1..560
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL039h11_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO₂"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"

BASE COUNT 92 a 186 c 173 g 109 t

Query Match 20.9%; Score 443.8; DB 10; Length 560;
Best Local Similarity 99.6%; Pred. No. 6.9e-67;
Matches 445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGTGCTTACACAGCGCCCGACGCGCGTCCCTATCGTCATCAATGCCGCG 60
DB 114 ATGGCTGTGCTTACACAGCGCCCGACGCGCGTCCCTATCGTCATCAATGCCGCG 173
QY 61 TCGTTCCGGTCAAGAACGCGCAACAGCTGCTCGTGAGCTTCTCGTGGCTCCGCA 120
DB 174 TCGTTCCGGTCAAGAACGCGCAACAGCTGCTCGTGAGCTTCTCGTGGCTCCGCA 233
QY 121 CGCAAGTCCACCTCGCGCTCGGCTGTTACTGTGCGACCTGTTGGCGCTGGAC 180
DB 234 CGCAAGTCCACCTCGCGCTCGGCTGTTACTGTGCGACCTGTTGGCGCTGGAC 293
QY 181 ATCGTATGTTGCTCTGAGGTCGCGCTTGTTCACAGCGCGCGCTGGCGGATGTG 240
DB 294 ATCGTATGTTGCTCTGAGGTCGCGCTTGTTCACAGCGCGCGCTGGCGGATGTG 353
QY 241 ACTGTGGCTGCTATTGAGCTGTGTCACAGCGCGCGCTGCGTCAATGACCATTTGCCCT 300
DB 354 ACTGTGGCTGCTATTGAGCTGTGTCACAGCGCGCGCTGCGTCAATGACCATTTGCCCT 413
QY 301 CGTACGACAGTACGCTGACGCTGGGACACCTCGTGTGCGACATCATGGCGGAG 360
DB 414 CGTACGACAGTACGCTGACGCTGGGACACCTCGTGTGCGACATCATGGCGGAG 473
QY 361 AAGTCCGCTTACCTCCACTCCATCAAGAGGGGCTGACCGGCTGTGGATTGACCACTCC 420
DB 474 AAGTCCGCTTACCTCCACTCCATCAAGAGGGGCTGACCGGCTGTGGATTGACCACTCC 533
QY 421 TGGTCTCGCCCAAGGCTGTGGGCAAG 447
DB 534 TGGTCTCGCCCAAGGCTGTGGGCAAG 560

RESULT 12
AV629117 436 bp mRNA linear EST 15-DEC-2000
LOCUS AV629117 Chlamydomonas reinhardtii 5% to 0.04% CO₂ Chlamydomonas
DEFINITION reinhardtii cDNA clone LCL052d05_r 5', mRNA sequence.
ACCESSION AV629117
VERSION AV629117.1 GI:10791751
KEYWORDS Chlamydomonas reinhardtii.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 436)

AUTHORS Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1..436
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="LCL052d05_r"
/clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO₂"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"

BASE COUNT 66 a 145 c 138 g 87 t

Query Match 20.4%; Score 432.8; DB 10; Length 436;
Best Local Similarity 99.5%; Pred. No. 5.2e-65;
Matches 434; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 CGGCTGTACTGTGTCACCTGTCGCGTGGACATCGGCGTGGATGCTGCTGCTG 199
DB 1 CGGCTGTACTGTGTCACCTGTCGCGTGGACATCGGCGTGGATGCTGCTGCTG 60
QY 200 AGTGGCCCTTGTGTCACGCGCGCTGGCGATGTCGCTGCTGCTGCTGCTGCTG 259
DB 61 AGTGGCCCTTGTGTCACGCGCGCTGGCGATGTCGCTGCTGCTGCTGCTGCTG 120
QY 260 AGTGGTCAAGCGCGCGCTGTCATGACCATGTCGCTGCTGCTGCTGCTGCTGCTG 319
DB 121 AGTGGTCAAGCGCGCGCTGTCATGACCATGTCGCTGCTGCTGCTGCTGCTGCTG 180
QY 320 ACCTGTTGGACCTGCTGCTGTCGACATCATGGCGAGAGGTCGCTGCTGCTGCTG 379
DB 181 ACCTGTTGGACCTGCTGCTGTCGACATCATGGCGAGAGGTCGCTGCTGCTGCTG 240
QY 380 CCATCAAGAGGCGTGCACCGCTGTGATTGACCACTGCTGCTGCTGCTGCTGCTGCT 439
DB 241 CCATCAAGAGGCGTGCACCGCTGTGATTGACCACTGCTGCTGCTGCTGCTGCTGCT 300
QY 440 GGGCAAGACCGCTGTCACGCGCGCTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTG 499
DB 301 GGGCAAGACCGCTGTCACGCGCGCTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 500 ACAAGCGCTTGGCGCTGCTGCAAGCGCGCTATTGAGGTCGCGCGCTGCTGCTGCTG 559
DB 361 ACAAGCGCTTGGCGCTGCTGCAAGCGCGCTATTGAGGTCGCGCGCTGCTGCTGCTG 420
QY 560 GCCCGCGCGAGGACTG 575
DB 421 GCCCGCGCGAGGACTG 436

RESULT 13
BE024926 552 bp mRNA linear EST 06-JUN-2000
LOCUS BE024926
DEFINITION BE024926 C. reinhardtii CC-1690, normalized, Lambda zap II
ACCESSION BE024926
VERSION BE024926.1 GI:8287367
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 552)
AUTHORS Grossman,A., Davies,J., Pederspiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Sillflow,C., Stern,D. and Surzycki,R.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
JOURNAL Unpublished (2000)
COMMENT Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.

FEATURES
source Location/Qualifiers
1..552
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II".
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT 91 a 185 c 167 g 109 t
ORIGIN
Query Match 19.9%; Score 422.8; DB 10; Length 552;
Best Local Similarity 99.5%; Pred. No. 2.8e-63;
Matches 424; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTGTGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATGCGCG 60
Db 127 ATGGCTGTGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATGCGCG 186
QY 61 TCGTTGCGGTGTCAGAAGACCGGCAACAGCTGCTGCGTGAGCTTGTCTGCGCTCGCA 120
Db 187 TCGTTGCGGTGTCAGAAGACCGGCAACAGCTGCTGCGTGAGCTTGTCTGCGCTCGCA 246
QY 121 CGCAAGTCCACCTCGCGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 180
Db 247 CGCAAGTCCACCTCGCGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 306
QY 181 ATCGGTGATGTTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 240
Db 307 ATCGGTGATGTTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 366
QY 241 ACTGTTGCGCTGCTTATGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 300
Db 367 ACTGTTGCGCTGCTTATGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 426
QY 301 CGCTACGACCACTAGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 360
Db 427 CGCTACGACCACTAGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 486
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGTCACCGCGTGTGATGACCAACCC 420
Db 487 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGTCACCGCGTGTGATGACCAACCC 546
QY 421 TGGTTC 426
|||||

Db 547 TGGTTC 552

RESULT 14
AV641583
LOCUS
DEFINITION AV641583 Chlamydomonas reinhardtii 5' mRNA linear EST 15-DEC-2000
CDNA clone HCL036g08_r 5', mRNA sequence.
ACCESSION AV641583
VERSION AV641583.1 GI:10784911
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 517)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source Location/Qualifiers
1..517
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="HCL036g08_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT 82 a 176 c 159 g 100 t
ORIGIN
Query Match 19.8%; Score 421.2; DB 10; Length 517;
Best Local Similarity 99.3%; Pred. No. 5.3e-63;
Matches 423; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCTGTGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATGCGCG 60
Db 92 ATGGCTGTGCTCTACCAAGCGCGCCAGCGCGCTCTATCGTCAATGCGCG 151
QY 61 TCGTTGCGGTGTCAGAAGACCGGCAACAGCTGCTGCGTGAGCTTGTCTGCGCTCGCA 120
Db 152 TCGTTGCGGTGTCAGAAGACCGGCAACAGCTGCTGCGTGAGCTTGTCTGCGCTCGCA 211
QY 121 CGCAAGTCCACCTCGCGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 180
Db 212 CGCAAGTCCACCTCGCGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 271
QY 181 ATCGGTGATGTTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 240
Db 272 ATCGGTGATGTTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 331
QY 241 ACTGTTGCGCTGCTTATGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 300
Db 332 ACTGTTGCGCTGCTTATGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 391
QY 301 CGCTACGACCACTAGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 360
Db 392 CGCTACGACCACTAGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 451
QY 361 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGTCACCGCGTGTGATGACCAACCC 420
Db 452 AAGTCCGCTACTTCCACTCCATCAAGAGGCGGTCACCGCGTGTGATGACCAACCC 511
|||||

DB 473 AAGGTCGGCTACTTCCACTTCATCAAGAAAGCGGTCCACCGCTGGGATTGACCAACCC 532

Search completed: June 3, 2003, 15:27:11

Job time : 2082.3 secs

QY 421 TGGTTC 426

Db 512 TGGTTC 517

RESULT 15

AV641989 532 bp mRNA linear EST 15-DEC-2000

LOCUS AV641989 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii

DEFINITION cDNA clone HCL043e06_r 5', mRNA sequence.

ACCESSION AV641989

VERSION AV641989.1 GI:10785317

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 532)

Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y. and Tabata, S.

Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii

DNA Res. 7 (5), 305-307 (2000)

20539644

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

Location/Qualifiers

1..532

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db_xref="taxon:3055"

/clone="HCL043e06_r"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"

BASE COUNT 89 a 179 c 162 g 102 t

ORIGIN

Query Match 19.5%; Score 415.2; DB 10; Length 532;

Best Local Similarity 99.3%; Pred. No. 5.8e-62;

Matches 417; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCTGTTCCTCTACCGCCGCCCGCCAGCAGCGCGGTCTATCGTCATCAATGCCGCG 60

Db 113 ATGGCTGTTCCTCTACCGCCGCCCGCCAGCAGCGCGGTCTATCGTCATCAATGCCGCG 172

QY 61 TCCTTCGGTCTCAAGAGACCGGACACAGCTGCTGGTGAGCTTGGTCTGGCTCGGCA 120

Db 173 TCCTTCGGTCTCAAGAGACCGGACACAGCTGCTGGTGAGCTTGGTCTGGCTCGGCA 232

QY 121 CGCAAGTCCACCTCGCGCTCGGTGTTACTGGTGCCACTGGTGCCACTGGCGCTGGAC 180

Db 233 CGCAAGTCCACCTCGCGCTCGGTGTTACTGGTGCCACTGGTGCCACTGGCGCTGGAC 292

QY 181 ATCGTGATGTTGCTGTGAGGTGCGCCCTTGGTCCAGACGGCGGCTGGCGGATGTG 240

Db 293 ATCGTGATGTTGCTGTGAGGTGCGCCCTTGGTCCAGACGGCGGCTGGCGGATGTG 352

QY 241 ACTGGTGGCTGCTATTGAGCTGTCAAGCGCGGCCACCGCGTCATGACCATTCGCCCT 300

Db 353 ACTGGTGGCTGCTATTGAGCTGTCAAGCGCGGCCACCGCGTCATGACCATTCGCCCT 412

QY 301 CGGTACGACCACTAGCTGACGCTGGGACACCTTCGGTGGTTCGGGACATCATGGCGAG 360

Db 413 CGGTACGACCACTAGCTGACGCTGGGACACCTTCGGTGGTTCGGGACATCATGGCGAG 472

QY 361 AAGTTCGGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATTGACCAACCC 420

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:49:44 ; Search time 51.0928 Seconds
(without alignments)
12748.991 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atggctgttgcctaccag.....ccgccaccgcagtcgcc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/6CTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478	22.5	2267	4	US-08-679-645-25
2	447.6	21.1	2542	3	US-08-941-445A-6
3	281.2	13.2	2067	4	US-09-388-743-21
4	266.6	12.6	2274	4	US-09-388-743-17
5	259.6	12.2	2176	4	US-09-388-743-13
6	255.4	12.0	2202	4	US-09-388-743-1
7	219.2	10.3	2097	3	US-08-941-445A-10
8	218.2	10.3	2825	4	US-09-196-390-5
9	213.4	10.0	2380	1	US-08-572-951-3
10	207	9.7	2248	4	US-09-345-214-20
11	206.6	9.7	1798	4	US-09-345-214-16
12	206.6	9.7	2019	4	US-09-345-214-15
13	161	7.6	2007	3	US-08-941-445A-8
14	161	7.6	2085	1	US-08-572-951-2
15	118.2	5.6	2239	4	US-09-196-390-1
16	109.6	5.2	2348	4	US-09-388-743-5
17	108	5.1	1758	3	US-08-836-567-3
18	106.8	5.0	2277	1	US-08-676-967-5
19	106.8	5.0	2277	1	US-08-676-974-5
20	106.8	5.0	2277	2	US-09-098-487-5
21	106.8	5.0	2418	4	US-09-388-743-25
22	106.6	5.0	2360	3	US-08-836-567-9
23	102.6	4.8	1528	4	US-09-345-214-6
24	102.6	4.8	1620	3	US-08-941-445A-20
25	102.6	4.8	1752	3	US-08-941-445A-12
26	102.6	4.8	2008	4	US-09-345-214-12
27	102.6	4.8	2491	4	US-09-345-214-5

Sequence 4, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 7, Appli
Sequence 182, App
Sequence 177, App
Sequence 10, Appl
Sequence 10, Appl
Sequence 5, Appli
Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-679-645-25
; Sequence 25, Application US/08679645
; Patent No. 6350934
; GENERAL INFORMATION:
; APPLICANT: Zwick, Michael G.
; APPLICANT: Edington, Brent E.
; APPLICANT: McSwiggen, James A.
; APPLICANT: Merlo, Patricia Ann Owens
; APPLICANT: Guo, Lining
; APPLICANT: Skokut, Thomas A.
; APPLICANT: Young, Scott A.
; APPLICANT: Folkerts, Otto
; APPLICANT: Merlo, Donald J.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; TITLE OF INVENTION: MODULATION OF GENE EXPRESSION
; TITLE OF INVENTION: IN PLANTS
; NUMBER OF SEQUENCES: 1263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/679,645
; FILING DATE: July 12, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,135
; FILING DATE: July 13, 1995
; APPLICATION NUMBER: 08/300,726
; FILING DATE: September 2, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 219/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
; LENGTH: 2267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-679-645-25

Query Match 22.5%; Score 478; DB 4; Length 2267;
Best Local Similarity 59.8%; Pred. No. 3.6e-72;
Matches 986; Conservative 0; Mismatches 563; Indels 99; Gaps 7;
QY 161 GTGCGACATTCGCGCTGGACATCGTGGTTCGCTGAGTGGCCCTTGTGTCACAGA 220
DB 403 GCGCCAGCGCGCATGAACGTCGCTTCGTCGCGCGCCGAGATGGCGCGTGGAGCAAGA 462
QY 221 CGGGGGCTGGCGCATGTGACTGTGGTGGCTGCCTATTTAGCTGTGTCAAGCGCGCCACC 280
DB 463 CGGGGGCTGGCGCATGTGACTGTGGTGGCTGCCTATTTAGCTGTGTCAAGCGCGCCACC 522
QY 281 GCGTCATGACCATTCGCGCTGACGACAGTACGCTGACGCGCTGGGACACCTCGGTGG 340
DB 523 GTGTGATGTCGTCCTCCCGCTACGACAGTACAGGAGCGCTGGGACACCGGCTCG 582
QY 341 TGTGTGACATP-----CATGGCGGAGAGTCCGCTACTTCCACTCCATCAAGA 388
DB 583 TGTCCGAGATCAAGATGGGAGACAGTACGACGCGTCAAGTTCCTTCCACTGCTACAGC 642
QY 389 AGGGGTGACCGCTGTGATGATGACACCGCTGTTCTGCTGCGCAAGTCTGGGCAAGA 448
DB 643 GCGGATGACCGCTGTGTTGATGACACCGCTGTTCTGCTGAGAGGTTTGGGAAAGA 702
QY 449 CCGGCTTCCAAAGTGTACGGCGCGCGCTCGGGGCTGACTACCTGGGACAAACCAAGCGT 508
DB 703 CCGAGGAGAGATCTACGGGCGCTGACGCTGGAAGGACTACAGGGACAAACCAAGC 762
QY 509 TGCCCTGTTTGCAGGCGCTATTGAGGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
DB 763 TCAGCCTGCTATGCCAGGAGCACTTTAAGTCTCAAGAGTCTTGAAGCTTCAACAAAC 822
QY 562 -----CCCGGAGGAGTGGTCTTCTGCTGCGCAAGCTGGCAAGCTGGCACTCCG 604
DB 823 CATACTTCCGGACCATACGGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 882
QY 605 CCGTGTGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664
DB 883 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 942
QY 665 AGTCGCTGCTGCTATCCACATCGCTTCCAGGCGCGCATGTGGGAGGAGCTTCA 724
DB 943 AGACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1002
QY 725 AGGACAGGAGTGTGCGCGCGCGCTTTCACAAAGCTGGCGCTTCGCGAGGCTATGCCA 784
DB 1003 CGGAGCTGAACCTCCGGAGAGATTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1062
QY 785 AGTTTACACTGAGGCGCGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 844
DB 1063 AG-----CCCGTGAAG----- 1074
QY 845 CCTACAGAGATCAACTGCTGAAGGTTGGCATTTATCGCGCGCGCAAGCTGGTACTG 904
DB 1075 ---GCGGAGATCAACTGATGAAGCGCGGATCTTCGAGGCGCGAGGCTCTTACCG 1131
QY 905 TGTGCGCAACTACGGGACCGAGATGCTGCGGATGCGCGCGCGGCTGAGCTGGAGTGACA 964
DB 1132 TCAGCCCTTACTAGCGGAGGAGCTCATCTCCGCGCATCGCGGCTGCGAGCTCGACA 1191
QY 965 CCGTCATCCGCGGCAAGGCAATTGAGGCGCATTTGACGCGCATGACATTCAGAGTGA 1024
DB 1192 ACATCATGCGCTTACCGGCGCATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1251
QY 1025 ACCCGAGACCGAAGTCTTCTGCTGCGCGCTACGACGAGAACAGCTTCTACGCGGCA 1084

DB 1252 ACCCCAGCAGGACAAAGTATATGCGCGTGAAGTACAGCTGTGCGAGCGCGGTGAGGCCA 1311
QY 1085 AGSGCGCGCGCAGGAGGCGCTGCGAGCGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1144
DB 1312 AGCGCTGAACAGGAGGCGCTGCGAGCGAGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 1371
QY 1145 CCGTGTTCGCTTCATCGCGCGCGCTGCGAGGAGCAGAGGCTGCGACATCACTCTGCGCG 1204
DB 1372 CGCTGTGCGCTTCATCGCGAGCTGGAAGAGCAGAGGACCCGACGCTCATGCGCGCG 1431
QY 1205 CCGTGCCTAAGATCCCTGCGCGCGCGCTGCGAGGAGCAGAGGCTGCGACATCACTCTGCGCG 1261
DB 1432 CCATCCCGAGCTCATGAGATGTTGAGGAGCAGTGCAGATGCTTCTGCTGCGCGCGCGCA 1491
QY 1262 AGCGCGCTTACGAGAGCTGGTGAACCGCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1321
DB 1492 AGAAGAGTTCAGCGCGCTCATGAGCGCGGAGGAGGAGTTCGCCAGGCAAGGTGCGCG 1551
QY 1322 GGTGTCAAGTTCCTGCGCGCGCGCTGCGGCGCATGCTCACCCTGCGCGCGCGCTGCTGCTGCTG 1381
DB 1552 CCGTGTCAAGTTCAGCGCGCGCTGCGGCGCATGCTGCGGCGCGCGCTGCTGCTGCTGCTGCTG 1611
QY 1382 TGTGCGCTGCGCTGCGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441
DB 1612 CCGTACCGCGCTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1671
QY 1442 CCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
DB 1672 CGCGCTGCGCTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1731
QY 1502 GCTTCCATGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1555
DB 1732 GCTTCCATGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1791
QY 1556 CCGTGCCTGCGCTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1615
DB 1792 AGTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1851
QY 1616 TGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1675
DB 1852 TGTGAGGAACTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1911
QY 1676 GCTGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1735
DB 1912 ACCTG 1968
QY 1736 TCAAGTGTGCGCTGCGGAGGAGTCCCG 1765
DB 1969 TCGCGCGCTGCGGAGGAGGAGTCCCG 1998

RESULT 2
US-08-941-445A-6
; Sequence 6, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:


```

; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Typha latifolia
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(1956)
US-09-388-743-17

Query Match      12.6%; Score 266.6; DB 4; Length 2274;
Best local Similarity 52.9%; Pred. No. 8.6e-37;
Matches 828; Conservative 0; Mismatches 639; Indels 98; Gaps

QY 172 GCCTGGACACVGTGATGTTGCTGTGAGGTGCGCCCTTGTTGTCACAGCGGCGCCTG 231
   || || || || || || || || || || || || || || || || || || || ||
Db 373 GGGATGAACCTAGTCTTTGTGTAGCTGAGATGGCTTCCATGGACGAAGACTGGAGCCTT 432
   || || || || || || || || || || || || || || || || || || || ||
QY 232 GGCATGTGACTGGTGCCTTCATTTAGCTGTGTCAACGGCGGCCACCGGTCATGACC 291
   || || || || || || || || || || || || || || || || || || || ||
Db 433 GGTGATGTTCTTGGAGGACTCCACC CGGCAATTGGCGGCAAAATGGACATCGAGTTAATGTT 492
   || || || || || || || || || || || || || || || || || || || ||
QY 292 APTGCCCCCTCGCTACGACCAAGTACGCTGACGGCTGGGACACACCTCGGTGGTCTGTCGACAT- 350
   || || || || || || || || || || || || || || || || || || || ||
Db 493 ATAGCGCCACGTTATGATCAATACATGATGCTTGGGATACAGATGCCTTGTTGAGTTG 552
   || || || || || || || || || || || || || || || || || || || ||
QY 351 -----CATGGCGGAGAAGTCCGCTACTTCCATCCATCAAGAAGGCGGTGCAC 399
   || || || || || || || || || || || || || || || || || || || ||
Db 553 AAAGTTGGGATAGGTTGAAACCGTCGCTCTTTTCACTGCTATAAAAGAGAGATTGAT 612
   || || || || || || || || || || || || || || || || || || || ||
QY 400 CGCTGTGGATTGACCAACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCCGGCTCCAAG 459
   || || || || || || || || || || || || || || || || || || || ||
Db 613 CGAGTTTGTGCAATCACCTATGTTTCTTGCGAAGGCTGGGGGAAAACGTGTGGGAAG 672
   || || || || || || || || || || || || || || || || || || || ||
QY 460 CTGTACGGCCCCCGCTCCGCGCGTACTACCTGGACAACCAAGCGGCTGCGCCCTGTTTC 519
   || || || || || || || || || || || || || || || || || || || ||
Db 673 ATTTATGGTCTTAACACTGGAACAGACTATCAGGACAATCAGCTACGCTTCAGCTTCTA 732
   || || || || || || || || || || || || || || || || || || || ||
QY 520 TGCAAAGCCGCTATTGAGCGTCCGCCGCTGCTGCCCTTCCGCTTCCGCTTCTCTCTCT 792
   || || || || || || || || || || || || || || || || || || || ||
Db 733 TGCAGGACGATTTGGAAGCTCTAGAAATCTAAATCTCAACAACAGTGAATCTTTCTCTCT 792
   || || || || || || || || || || || || || || || || || || || ||
QY 565 -----GGCGAGGACTGGCTTCGTGTGGCCAAAGCTGCACTCCGCCCTGGTGCC 615
   || || || || || || || || || || || || || || || || || || || ||
Db 793 GGTCCATTATGGGAAGATGTAATCTTCATTTGCAATGATTTGGCAGACTTCCTCTCGCA 852
   || || || || || || || || || || || || || || || || || || || ||
QY 616 GTCCGTGTGAAGCAGAGTACCAAGCCCAAGGCCAGTTTCAACAAGGCCAAAGTCGSGTGTG 675
   || || || || || || || || || || || || || || || || || || || ||
Db 853 TGCTACTTAAAGAGCATGTACCATCCCGTGGCAATTTACAAGAAGCCCAAGGTTGCTTTC 912
   || || || || || || || || || || || || || || || || || || || ||
QY 676 GCTATCCACAACTCCGCTTCCAGGCCCCCATGTGGGAGGAGGCTTTCAGAGCAACAGAG 735
   || || || || || || || || || || || || || || || || || || || ||
Db 913 TGCATTCACAATATATCATACCAAGGTCGATTTTCTCCCTCAGACTTCGAAATTTCTCAAT 972
   || || || || || || || || || || || || || || || || || || || ||
QY 736 CTGCCCCAGGCCCTTTTGACAAGCTGGCTTCTCGGACGGCTATGCCAAGGTTTACACT 795
   || || || || || || || || || || || || || || || || || || || ||
Db 973 CTTCGGAG-----AATTCAATCT 993
   || || || || || || || || || || || || || || || || || || || ||
QY 796 GAGGCCACCCCATGAGGAGGACGAGAGCCCGCTGAGGGAAAGACCTACAAGAAG 855
   || || || || || || || || || || || || || || || || || || || ||
Db 994 TCTTTTCAGTTTTCATTATGGGTACAACAAGC-----TGTAAGGGATGAAG 1041
   || || || || || || || || || || || || || || || || || || || ||
QY 856 ATCAACTGGCTGAAGGGTGGCAATTATGCCCCCGACAGCTGGTGACTGTGCGCCCAAC 915
   || || || || || || || || || || || || || || || || || || || ||
Db 1042 ATAAATTGGATGAAGGACGAAATCTTAGAATCAGACAGGGTGTTTTACAGTAGTGCATAT 1101
   || || || || || || || || || || || || || || || || || || || ||
QY 916 TAGCGACCGAGATCGCTCCGATCGCGCGCGGTGTGGAGCTGGACACCGTCAATCCGC 975
   || || || || || || || || || || || || || || || || || || || ||
Db 1102 TATGCACAGAGCTCTTTTCAGSAGAAAGAGGGAGTCGAGTTGGACAACATTTTCGGT 1161
   || || || || || || || || || || || || || || || || || || || ||
QY 976 GCGAAGGGCATTTGAGGGCATTTGAAACGSCATGGACATTGAGGAGTGGAAACCCCAAGACC 1035
   || || || || || || || || || || || || || || || || || || || ||
Db 1162 GTGACCAAGTATCACAGGAATTGTGAATGGAATGATGTTAATGAGTGAATCCATTACA 1221
   || || || || || || || || || || || || || || || || || || || ||
QY 1036 GACAAGTTCTGTCTCGCCCTACGACCCAGAACAGCGTCTACGCCGCAAGGCCGCGCC 1095
   || || || || || || || || || || || || || || || || || || || ||

```

Matches 795; Conservative 0; Mismatches 649; Indels 96; Gaps 5;

QY 176 TGGACATCGTGATGTTGCTGAGTGGCGCCCTTGGTCCAAAGACGGCGCGCTGGCG 235
Db 349 TGAACITGGTGTGTTGTTGAGTGTAGTCCATGGAGCAAACTGGGGCGCTTGGCG 408
QY 236 ATGTGACTGGTGGCTTATGAGCTGTCAAGCGCGGCCACCGCTCATGACCATTTG 295
Db 409 ATGTTCTTAGAGGATTGCCACCTGCTATGGCTCAATTTGGGCACAGGGTCATGACCGTGG 468
QY 296 CCGCTCGCTACGACCACTAGCTAGCGCTGGGACACCTCGGTGGTTCGTGGACAT----- 350
Db 469 TCCACGATATGACCAATAAAGATATCTGGGATCAAACTGTCACAGTTGAGTTAAAG 528
QY 351 -----CATGGCGAGAGGTCGCTACTTCCACCTCCAAAGAAAGGGCGTCCACCGG 403
Db 529 TTGGGGATGATTTGAACCTGTCGGTCTTCCACTGCTACAAAAGGGGATGATCGGG 598
QY 404 TGTGATTGACCAACCCCTGTTCTTCCCAAGGTCTGGGCAAGACCGGCTCCAAAGCTGT 463
Db 589 TTTTGTGGATCACCTATGTTTTCGAGAAGGTTTGGGGGAAAACAGAGGAAATAT 648
QY 464 ACGGCCCCCGCTCCGGCTGACTACTGGACAAACACAGCCCTTCGCCCTGTTGCA 523
Db 649 ATGGTCTCTTACAGAACAGATATTCACAGACAATCAACTAAGATTACGCTTTTGTGC 708
QY 524 AGCCCGCTATTGAGGCTGCCCGGTGCTGCCCTTCGGCCCGCGGAGGACTGC----- 576
Db 709 TGCACCTCTGGAAGTCTCAAGACTTCTAATCTCAACACACAACTATCTCTGGAC 768
QY 577 -----GTCTTCGTGGCCCAACGACTGGCACTCCGCCCTGGTGGCGCGTCC 619
Db 769 CATATGAGATGATGTTGTTTATTTGCAAGGATTTGGCTTCTGCTACTGCGCTGCT 828
QY 620 TGCCTGAGGAGCTAGTACCACCAAGGCGAGTTCACCAAGGCCAAGTCCGTTGCTGCTA 679
Db 829 ACTTTGAAACTATGTACCAATACATGTTATTTACATGAATGTAAGTTGTCATTTTGA 888
QY 680 TCCACAACTCGCTTCCAGGGCGGATGTGGAGGAGGCTTTCAGGACACCAAGCTGC 739
Db 889 TTTATATATTTGCTTACAGGCGGATTTGCTTTTCGGACTTTGAACCTCTTAATCTCC 948
QY 740 CCGAGGCGGCTTTGACAAGCTGGCTTTCGGAAGGCTATGCCAAGGTTTACACTAGG 799
Db 949 CCAATAAATTTAAATCTTCAATTTGATTTCAATGGAATGATGACAAACCT----- 998
QY 800 CCACCCCATGGAGGAGGACGAGGAGGCGGCTGACGGGAAGACCTCAAGAAGATCA 859
Db 999 -----GTCAAGGAGGAAATAA 1017
QY 860 ACTGGCTGAAGGTTGCTATTCGCGCGCCGACAAAGCTGTGCTGTCGCCCACTACG 919
Db 1018 ATTGGATGAGGCTGAATAATAGATGTATAGGTTGCTTGACCGTGGAGCCATATTATG 1077
QY 920 CGACCGAGATCGCTGCGGATGCGCGGCTGTGGAGCTGGACACCGTCAATCCGCCCA 979
Db 1078 CCAAGAGCTGTCTCAGGCTGAGAAAGGCTGTGAGTTGGGCAATATCTCGCATGA 1137
QY 980 AGGCAATTGAGGCTATGTAACGGCATGGACATTTGAGGAGTGAACCCCAAGACCCACA 1039
Db 1138 AAACATCTGTGGAATAGTAATGGGATGACACCCAGGAGTGAATCCATTACAGACA 1197
QY 1040 AGTTCCTGTCTGCGCCCTACGACAGAACACGCTCTACCGCGGCAAGCGCGCGCCCAAG 1099
Db 1198 AATATATTCTACAACTAGCATCAACAACCTGATTGATGCAAAAACCTCTCTGTAAG 1257
QY 1100 AGCCCTCGAGGCGAGCTGGCTGCTGTGACCCACCGCCCGCCCTGTCGCCCTCA 1159
Db 1258 AAGCTTTGCAAGCTGAGTGTGGGCTGCTGTTAAACAAAACAAAGCTTGTGTCGCCCTTG 1317
QY 1160 TCGGCCCTCGGAGGACGAAGGCTGTGGACATCATCTTCGCCCGCGCTGCCCAAGATCC 1219
Db 1318 TTGGAAGACTAGATGACGAAAGGCTCAGACATTTCTAGCTGCAGCAATTCAGAACTTC 1377

QY 1220 TGCCCAACCCCAAGTGTGAGATGCCATCTTGGTTACCGCAAGCGCGCTACGAGAAGC 1279
Db 1378 T---TTTGAGAAATGTTCAAGTGATAGTACTTGGCACTTGCAGAAAGAAAGTTGGAGAGTG 1434
QY 1280 TGTGAAAGCCATCGGCACCAAGTACAAGGCGCGCCCAAGGGGTGGTCAAGTTCTCGG 1339
Db 1435 AACTTACATTACTTTAGGAAATGTTTCCAGACAAATTCAGAGCAATCTCAAAATCAAG 1494
QY 1340 CGCCCTGGCGGCACATGCTCACCGCGCGCGGCTCATCTGCTGGTGGCTCGCGTTCG 1399
Db 1495 TTCTTTAGTCTAGCAATCATGCGAGGAGCTGATATCTTGTATTCAGAGCAGATTCG 1554
QY 1400 AGCCCTGCGGCCCTGATCCAGCTGCAGCCATGACATACGTTACGTTCCGCGTGGTACCT 1459
Db 1555 AACCCTGTGCGCTCATTCAGCTTCAGGCCATGCGATATGGAATCTCTCCCTATGTAGCA 1614
QY 1460 CACCGCGCGCTTGGTGCACACCTCAAGGAGGCGCTCACCGGCTTCACATGGCGCC- 1518
Db 1615 CCACTGTGAGCTTGTGACACTGTCAAGAAGGCTTCACTGGCTTCCATATGGGCCCT 1674
QY 1519 -----CTGAACCCCGCAAGCTGGACGAGCTGACGCGGACGCGCTGGCGGCCACCGTGC 1573
Db 1675 TCAGTGTGGAGTGTGATGCGGTAGACAAAGCTGATGTACAAAGATTGTGAAACACGA 1734
QY 1574 GCGTGTCCAGCGAGGTGTTTGGCGGCGCGCTACCCGAGAGTGTGGCCAACTGCATCA 1633
Db 1735 AAAGGCGCCCTCAAAGTCTATGTAACACCTGCTTTTGTGGAGATGATCAAGAAGTGCATGA 1794
QY 1634 GCAGAGACCTGTCTCTGTCTCCAAAGCCCGCCAGAGTGGGA 1673
Db 1795 ACCAAGATCTCTATGGAAGGACCTGCAAGAAGTGGGA 1834

RESULT 6
US-09-388-743-1
; Sequence 1, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; FILE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2202
; TYPE: DNA
; ORGANISM: Curcuma zedoaria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)....(1974)
US-09-388-743-1

Query Match 12.0%; Score 255.4; DB 4; Length 2202;
Best Local Similarity 51.8%; Pred. No. 6.5e-35;
Matches 799; Conservative 0; Mismatches 646; Indels 96; Gaps 6;

QY 176 TGGACATCGTGATGTTGCTGAGTGGCGCCCTTGGTCCAAAGACGGCGCGCTGGCG 235
Db 395 TGAACITGATCTTGTGTTGAGTGTAGTCCATGGAGCAAACTGGAGGCTTGGTG 454
QY 236 ATGTGACTGGTGGCTTATGAGCTGTCAAGCGCGGCCACCGCTCATGACCATTTG 295
Db 455 ATGTCTTGGAGGTTTCCACCGCCCATGGCGCAAGGACACAGGCTGATGACTATAG 514
QY 296 CCGCTCGCTACGACCACTAGCTAGCGCTGGGACACCTCGGTGGTCTGTGACATCAGG 355
Db 515 CACCGCGACATGACCAATACAAAGATGATGGGATACGGTGTCTTTGTCGATTGAAG 574

QY	144	TGTTACTGGTCCCACTGGTGCCACTTCGCGCTGGACATCGTGAITGTGCTGCTGAGGI	203
Dd	588	TGAACCTTGGCCCTTTTGGCTGGGCCCTAATGTGATGAACGTCTGCTGGTGGCTTCIGRATG	647
QY	204	CGCCCTTGGTCCAAAGACGGCGGCCTTGGCGATGTGACTGGTGGCCCTGCCTATTGAGCT	263
Dd	648	TGCTCCCTTCGCAAGACAGTGGCCTTGGNATGTCGTGGGTGCTTTGCCTAAAGGCTCT	707
QY	264	GGTCAAGGGGGCCACC CGCGTCAAGCATTCGCCCCCTGCTGACACACAGTAGCTGAGCG	323
Dd	708	GGCAGGAGAGGACACCGTGTATTGGTCTGATACCRAGATATGAGAGTAGTCGCGAAGC	767
QY	324	CTGGGACACCTCGGTGGTGGACATCATGGCGAGAAGTCCGCTACTTCCCACCTCCAT	383
Dd	768	CCGGG--ATTTAGGTGTAAAGGAGAGCTTACAAGGTAGCTGGACAGGATTCAGAGTTPACT	825
QY	384	CAAGAAGGGCGTGACCGCGTGTGGATGTACCAACCCCTGGTTCCTGGCCAAAGTCTGGGG	443
Dd	826	TATTTTCACTTTACATTAUTGATGGAGTTGATTTTGTATTTCGTAGAAGCCCTCCCTTCGGG	885
QY	444	CAAAGCCGGCTCCAAGCTGTAGCGCCCGCGCTCGCGCGCTCAC TACTGGACAACCAAA	503
Dd	886	CAC-----CGGCACAATAATTATGGGGGAGAAAGATTGGATATTTTGAA	932
QY	504	CGCGTTCGCCCTGTTCTCAAGGGCGCTATTGAGGCTCCCGCGCTGCTGCCCTTCGCGCC	563
Dd	933	CGCATGATTTGTCTCAAGGGCGCTGTGAGGTTCCATGGTATGCTCCATGTGGCGG	992
QY	564	-----CGCGAGGACTGGTCTTCGTGGCCAAACNACTGGCACTCCGCCCTGGT	611
Dd	993	TACTGTCTATGGTGA TGCACTAGTTTTTCAATCTAATGATTGGCATACGCACTTCT	1052
QY	612	GCGCGTCTGCTGAAGGACGAGTACCAAGCCAAAGGCCAGTTTCA CCAAGGCCAACGTGCT	671
Dd	1053	GCCTGTCTATCTAAGGCCATATTACGGGACAAATGGTTTGATGCAGTATGCTCGCTCCT	1112
QY	672	GCTGCTATCCACAACATCGCCTTCCAGGGCCGATGTGGGAGGAGCTTTCAAGSACAC	731
Dd	1113	GCTTGTGATACACAACATTGCTCATCAGGCTCGTGGCCCTGTGAGACGACTTCGTCAATTT	1172
QY	732	GAAGTCGCCAGGGCGCCTTTGACAAGCTGSCCTTCTCGGACGCTATGCCAAGGTTTA	791
Dd	1173	TGACTTGGCTGAACACTACATCGACCACTTTCAAATGTATGACAACTTGTGGGGATCA	1232
QY	792	CACTCAGGCCCCCATGGGAGGAGGAGAGAGCCCGCTGACGGGAAGACCTACAA	851
Dd	1233	C-----	1233
QY	852	GAAGATCAACTGGCTGAAGGTGGCATTTATCGCCGCGACACAGCTGGTGACTGTGTCGC	911
Dd	1234	----ASCAACGTTTTTGTCTGGCGGCTGAAGACGCGACAGCGGGTGGTACCGCTTAGCAA	1289
QY	912	CAACTACGCGACGAGATCGCTGCCGATGCCCGCGGCTGTGGAGCTGCGACACCGCTCAT	971
Dd	1290	TGGCTACATGTGGAGCTGAACATTCGGAAGCGGGTGGGGCTCCACGACATCATARA	1349
QY	972	CC---GCGCCCAAGGCAATTGAGGCGATTTGTAAACGSCATGGACATTTGAGAGTGGAAACC	1028
Dd	1350	CCAGAACGACTCGAAGCTCGAGGGATCTGTGAACGSCATCGACATGAGCGAGTGGAAACC	1409
QY	1029	CAAGACCGACAAGTTCTCTCTCGGCCCTTACGACCCAGAACAGC-----GTCFA	1076
Dd	1410	CGCTGTGGACGTGCACTTCCACTCCGACCACTATACCAACTACAGTTGAGACGCTGGA	1469
QY	1077	CGCGCGCAAGGGCGCGCAAGAGGCGCCTGCAAGCGCGAGCTGGGCTTCCCTTGGAACCC	1136
Dd	1470	CACCGCAAGGSGCACTGCAAGSGCGCCTTGCAGCGGCGAGCTGGGCTCGAGTCCGCGA	1529
QY	1137	CACCGCCCGCTTGTTCGCTTTCATCGGCCCTTGWGAGGAGCAGAAAGGTTGTGCATCAT	1196
Dd	1530	CGACGTGCCCATGTCCGGTTTATCGGGGGCTTGSACACCAAGGGGCTGGACATCAT	1589

QY	1197	CCTGCCGCCCTGCCCCAAGATCCTGTGCCACACCCCAAGGTGCAGATCCCATCTCTGGTAC	1251
Db	1590	CGCCACACGATCCACTGTGATC---GCGGGGAGGACGTGTGACCTCGTATGCTGGGCAC	1646
QY	1257	CGGCAAGGCGCGCTACGAGAAGCTGGTGAACGCCATCGGCACCAAGTACAAGGGCCGCGC	1316
Db	1647	CGGGCGGGCCGACCTGGAGGACATGCTGCGGGGTTCGAGTCGGAGCACACGACAAAGT	1706
QY	1317	CAAGGGCGGTGTCAGATTCTCGGGGCCCTTGGGGGCACATGCTCACCGCGGGGCCGACATT	1376
Db	1707	GCGCGCGTGGTGGGGTTCTCGGTGCCCTTGGCGCACCGCATCACGCGGGGCGCGGACAT	1766
QY	1377	CATGCTGTGTGCCCTTGCSCGCTTGAGCCCTTGAGCCCTGCGGCCATGCCAGCTGACGCCCATGCAC	1436
Db	1767	CTGTCTGATGCCGTGCGGGTTTGAGCGCTGCGGGCTGAACAGCTCTACGCCATGGCGTA	1826
QY	1437	CGGTACCGTGCOCGTGGTGTAGCTTCCACCGGCGGCCCTGGTGCAGACCGTCAAGAGAGCGGT	1496
Db	1827	CGGGACCGTGCOCGTGGTGCACGCCCTGGGGGGCTCCGSGGACACGGTGGCGCGCTTCGA	1886
QY	1497	CACCGCTTCCACATGGGCGCCCTGAACCCGACAAAGCTGGACGAGGCTGACGCCGACGC	1556
Db	1887	CCGTTTCAACGACACCGGCTCGGTGGAGCTTCGACCGCGGAGGCGAACCAGATGAT	1946
QY	1557	CCTGGCGGCACCGTGCOCGTGCCAGCGAGGTGTTTGGGGGCGGCCCTACCCCGAGAT	1616
Db	1947	CGACGCGCTCTGCACCTGCTCAACACAGTACCCGGAATACAAGGAGAGCTGGGGCGCGCTG	2006
QY	1617	GGTGCCAACTGCATCAGCCAGGACCTGTCTCTGGTCCAAAGCCCGCCACAGAGTGGGAGGG	1676
Db	2007	CAGGGCGCGGGCATGGCCGAGAGACTCACTGGGACCACCGCGCGCTGTATGAGGA	2066
QY	1677	CTGTGAGGAGGAGGTGGTGTAC	1698
Db	2067	GTGTGCTGTCAAGCGCAAGTAC	2088

RESULT 8
US-09-196-390-5
; Sequence 5, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Luticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 21 588.9
; FILING DATE: 29-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 36 917.7

APPLICANT: KEELING, PETER L.
 APPLICANT: KNIGHT, HARRY E.
 APPLICANT: GUAN, MING
 TITLE OF INVENTION: MODIFICATION OF STARCH
 TITLE OF INVENTION: SYNTHESIS IN PLANTS
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN DABRY & CUSHMAN
 ADDRESSEE: Intellectual Property Group of
 ADDRESSEE: Pillsbury Madison & Sutro LLP
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/572.951
 FILING DATE: 15-DEC-1995

Query Match	10.0%;	Score 213.4;	DB 1;	Length 2380;
Best Local Similarity	32.3%;	Pred. No. 6.9e-28;		
Matches 542;	Conservative 239;	Mismatches 810;	Indels 86;	Gaps 7;

QY	149	CTGGTCCCACTGCTGCCACTTGCGCGCTGGACATCGTGAAGTTGCTGTGAGTCGGCC	208
		: :	
Dd	650	CNGNCGCNTYNGCNGGNCCNAAYGTNATAAGYTNTGNTNGCNWNGARTGYGCNC	709
		: :	
QY	209	CTTGTGCTCAAAGACGGCGGCGCTGGCGAATGTGACTGGCTGCGCTATTGACGTGGTCA	268
		: :	
Dd	710	CNTTYTYARACNCGNGNYTNGNGAYGTNTNGCNGCYNTCCNAARGCNYTNGCNK	769
		: :	
QY	269	AGCGGGCCACCGGGTCATGACCATTGCGCGCTCGCTACGACCATGTAAGTACGCTGGG	328
		: :	
Dd	770	GNGNGGNCAYMNGNTNATGGTNTATHCCNMGNFYAGNGARTAYGCGNARGCMNGS	829
		: :	
QY	329	ACAACCTCGGVTGCTGGACATCATGGGGGAGAAGTCGCGTACTTCCCACCTCCATCAAGA	388
		: :	
Dd	830	--AYIYINGNGTNMGNMGNNTAYAARGTNGCNGNCAR--GAYWSNGARTNACNTAYT	886
		: :	
QY	389	AGGCGGTGCACCCGGTGTGGATTGACCAACCCCTGGTTCCTGGCCAAGTCTGGGGCAAGA	448
		: :	
Dd	887	TYCAYSNTAYATHGAYGGNTNGAYTTYGTNTYTGARGCNCNCNTYTMGNCAFM	946
		: :	
QY	449	CCGGCTCCAAAGCTGTACGGCCCCCGCTCGGGCGCTGACTACCTGGAACAACCAACAGCGT	508

Db	738	CCCCAAGGTGGACGTGCACCTCGGGTCGGACGGGTACACCAACTACTTCCCTCGAGACACT	679
QY	1074	CTACGCCGCCAAGCGCGCGCCAAAGAGGGCCCTCAGCGCCGAGCTGGGCGTGCCTGTGGA	1133
Db	678	CGACGCTGGAAGCGGCAGTGC AAGCGGCCCTCAGCGGAGCTGGCGCTGGAATGCG	619
QY	1134	CCCCACGCCGCCCTGTTCGCGCTTCATCGCGCCCTGGAGGACGAAGGGTGTGCACAT	1193
Db	618	CGACGACGTCGCGCTGCTGGGTTTCATCGCGCGTCTGGATGCAG AAGGGCGCTGGACAT	559
QY	1194	CATCTGCGCGCCCTGGCCCAAGATCCTGGCCACCCCAAGGTGCAGATCCCATCTCGG	1253
Db	558	CATCGGGAGCGATGCCGTGATC---CGCGGGCAGGACGTGCAGCTGGTGTGCTGG	502
QY	1254	TACCGGCAAGCGCCGCTACGAGAAGCTGGTGAAGCCCATCGGCACCAAGTAC AAGGGCCG	1313
Db	501	CACCGGCGCGCGACGCTGGAACGAATGCTGCAGCACTTGGAGCGGAGCATCCCAACA	442
QY	1314	CGCCAAGGCGTGGTCAAGTCTTCGSGCGCCCTGGGCGACATGCTACCGCGCGCGCGGA	1373
Db	441	GGTGGCGGGTGGTGCGKTTCTCGTGCCTATGCGGCACTCGCATCACGGGGCGCGGA	382
QY	1374	CTTCATCTGTGTGCCCTCGCGTTTCAGGCCCTTGGGGCTGATCCAGTGCACGCCATGCA	1433
Db	381	CGTGCTGGTATGCCCTCGCGTTTCAGCCCTTGCGGCTGAACCAAGTCTACGCGATGCG	322
QY	1434	CTACGGTACCGTCCGCGTGTAGCTCCACCGCGGCGCTGGTGACACCGCTCAAGGAGGG	1493
Db	321	ATACGGCACGTCCTCTGTGGTGACGCGGTGGGCGGGCTCAGGACACACCTGGCGCGCTT	262
QY	1494	CGTCACCGGTTCCACATGGGCGCCCTCAACCCCGACAAGCTGGACGAGGTGACGCCGA	1553
Db	261	CGACCCGTTTCAGGACGCGCGGCTCGGTTGGACTTTTGACCGYCGCGAGGCCACAAGCT	202
QY	1554	CGCCCTGGCGCCACCGTGGCGGTGCACGAGGTGTTTGGGGCGCGCGCTACCCCGA	1613
Db	201	GATCAGGCGCTCAGGACCTGCTCGACACGTACCGGAACCTACGAGGAGAGCTGGAAG	142
QY	1614	GATGGTGGCAACTGCATCAGCCAGGACCTGCTCTGTCTCCAAAGCCCGCCCAAGTGGGA	1673
Db	141	TCTCCAGCGCGGCATGTCGAGCACTCAGCTGGGACCAACGCGGCTGAGCTCTACGA	82
QY	1674	GGGCGTCTGGAGGAGGTGGTGTAC	1698
Db	81	GGACGCTCTTGTCAAGGCCAAGTAC	57

Qy	1554	CGCCTCGCCGCCACCGTGGCGTGCCAGCAGAGTGTTTGGCGGCGCGCTACCCCGA	1613
Db	1604	GATCGAGGCGCTCAGGCACTGCTTCACACAGTACCGGAACCTACAGAGAGCTGGAAAG	1663
Qy	1614	GATGGTGGCCAACTGCATCAGCCAGACAGCTCTCTTCGAAGCCCGCCAGAAAGTGGGA	1673
Db	1664	TCTCCAGCGCGGGCATGTCGAGAGACCTCAGCTGGGACACGCGGCTGAGCTCTACGA	1723
Qy	1674	GGCGCTGCTGGAGAGAGTGGTGTAC	1698
Db	1724	GGACGTCTCTTGTCAAGGCCAAGTAC	1748

```

: RESULT 13
: US-08-941-445A-8
: Sequence #, Application US/08941445A
: Patent No. 6107060
: GENERAL INFORMATION:
: APPLICANT: Keeling, Peter
: APPLICANT: Guan, Haining
: TITLE OF INVENTION: Starch Encapsulation
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan
: STREET: 5370 Manhattan Circle

```

```

: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Wisner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2007 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2007
: US-08-941-445A-8

```


Search completed: June 3, 2003, 15:30:36
Job time : 81.0928 secs

This Page Blank (uspr)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:49:59 ; Search time 192.814 Seconds
(without alignments)
15416.986 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atgctgttgcctaccag.....ccgccaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478	22.5	2267	9 US-09-961-077-25	Sequence 25, Appl
2	314.8	14.8	2399	9 US-10-138-075-1	Sequence 1, Appl
3	277.8	13.1	2179	9 US-10-138-075-3	Sequence 3, Appl
4	218.2	10.3	2825	10 US-09-952-677-5	Sequence 5, Appl
5	118.2	5.6	2239	10 US-09-952-677-1	Sequence 1, Appl
6	102	4.8	2712	10 US-09-748-033-4	Sequence 4, Appl
7	94.2	4.4	790	10 US-09-966-881-44	Sequence 44, Appl
8	86.6	4.1	1266	9 US-10-145-415-98	Sequence 98, Appl
9	84.2	4.0	2307	9 US-09-893-519A-87	Sequence 87, Appl
10	83.2	3.9	1427	10 US-09-974-300-757	Sequence 757, Appl
11	82.2	3.9	805	10 US-09-966-881-43	Sequence 43, Appl
12	81	3.8	12441	9 US-09-988-384B-3	Sequence 3, Appl
13	81	3.8	13613	9 US-09-860-846-3	Sequence 3, Appl
14	81	3.8	13613	9 US-09-836-821-3	Sequence 3, Appl
15	81	3.8	13613	10 US-09-861-289-3	Sequence 3, Appl
16	76.4	3.6	813	10 US-09-966-881-45	Sequence 45, Appl
17	76	3.6	1318	9 US-09-934-900-15	Sequence 15, Appl
18	75.4	3.5	446	10 US-09-864-761-20699	Sequence 20699, A
19	75.2	3.5	1896	9 US-10-124-880-15	Sequence 15, Appl

20	74.8	3.5	1407	10 US-09-815-242-4089	Sequence 4089, Ap
21	74.8	3.5	15872	9 US-09-860-846-1	Sequence 1, Appl
22	74.8	3.5	15872	9 US-09-988-384B-1	Sequence 1, Appl
23	74.8	3.5	15872	9 US-09-836-821-1	Sequence 1, Appl
24	74.8	3.5	15872	10 US-09-861-289-1	Sequence 1, Appl
25	74.6	3.5	1266	9 US-10-145-415-104	Sequence 104, App
26	74.2	3.5	1770	9 US-10-100-957A-1	Sequence 1, Appl
27	73.6	3.5	88421	9 US-09-976-059-1	Sequence 1, Appl
28	72.4	3.4	3468	9 US-09-988-462-2	Sequence 2, Appl
29	72.2	3.4	259	10 US-09-294-093B-3520	Sequence 3520, Ap
30	72.2	3.4	824	10 US-09-894-633A-83	Sequence 83, Appl
31	71.4	3.4	3624	9 US-09-988-462-6	Sequence 6, Appl
32	71.4	3.4	12733	9 US-10-032-393-47	Sequence 47, Appl
33	71.4	3.4	12739	9 US-10-032-393-8	Sequence 8, Appl
34	71.2	3.4	1182	9 US-09-364-847-1	Sequence 1, Appl
35	71.2	3.4	1926	9 US-09-364-847-9	Sequence 9, Appl
36	71.2	3.4	1926	9 US-09-364-847-18	Sequence 18, Appl
37	70.8	3.3	1415	9 US-09-934-900-11	Sequence 11, Appl
38	70.6	3.3	1635	10 US-09-864-761-20241	Sequence 20241, A
39	70.6	3.3	1973	10 US-09-864-761-3471	Sequence 3471, Ap
40	69.8	3.3	1075	10 US-09-864-761-19241	Sequence 19241, A
41	69.8	3.3	1403	10 US-09-864-761-2513	Sequence 2513, Ap
42	69.8	3.3	2010	12 US-10-032-717-9	Sequence 9, Appl
43	69.8	3.3	4889	9 US-09-860-846-34	Sequence 34, Appl
44	69.8	3.3	4889	9 US-09-988-384B-34	Sequence 34, Appl
45	69.8	3.3	4689	9 US-09-836-821-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1

US-09-961-077-25
; Sequence 25, Application US/09961077
; Publication No. US20030014775A1

GENERAL INFORMATION:

APPLICANT: Zwick, Michael G.
Edington, Brent E.
McSwiggen, James A.
Merlo, Patricia Ann Owens
Guo, Lining
Skokut, Thomas A.
Young, Scott A.
Folkerts, Otto
Merlo, Donald J.

TITLE OF INVENTION: COMPOSITION AND METHODS FOR
MODULATION OF GENE EXPRESSION
IN PLANTS

NUMBER OF SEQUENCES: 1263

CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/961,077

FILING DATE: 21-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/679,645

FILING DATE: July 12, 1996

APPLICATION NUMBER: 60/001,135

FILING DATE: July 13, 1995

APPLICATION NUMBER: 08/300,726

FILING DATE: September 2, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 219/247

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 2267 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-961-077-25

Query Match 22.5%; Score 478; DB 9; Length 2267;

Best Local Similarity 59.8%; Pred. No. 4.9e-100;

Matches	986;	Conservative	0;	Mismatches	565;	Indels	99;	Gaps	7;
---------	------	--------------	----	------------	------	--------	-----	------	----

Qy 161 GTGCCACTTGCGCGCTGGACATCGTGATGGTTGCTGCTGAGGTCGCCCCCTTGGTCCAAGA 220

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

.....

QY 221 CGGGCGCCTGGGCGATGTGACTGGTGGCCTGCCTATTGAGCTGGTCAAGCGCGGCCACC 280

Db 463 CCGCGGCCTCGGCGACGTCTCGGCGCCTCGCGCGCATGGCGGAATGGCACC 522

Downloaded from <http://ajphaphysocpharm.sagepub.com> at National Archive Publishing Co on June 11, 2015

Db 523 GTGTCATGGTCGTCTCTCCCCGCTACGACCAAGTACAAGGACGCCCTGGGACACCAGCGTCG 582

QY 341 TCGTGGACAT-----CATGGCGGAGAGGTCGCTACTTCCACTCCATCAAGA 388

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 389 AGGGCGTGCAACCGCGTGTGGATTGACCAACCCCTGGTTCCTGGCCAAGGTCCTGGGCAAGA 448

Db 643 GCGGAGTGGACCGCGTTCGTTGACCAACCCAC"GTTCCTGGAGAGGGTTTGGGAAAGA 702

1

Db 703 CCGAGGAGAAGATCTACGGGCTGACGCTGGAACGGACTACAGGGACAACCAGCTGCGGT 762

Qy. 509 TCGCCCTGTTCTGCAAGGCCGCTATTGAGGCTGCCCGCGTGCTGCCCTTCGGC----- 561

.....

QY 562 -----CCCGGCGAGGACTGCGTCTTCGTGGCCACGACTGGCACTCCG 604

Db 823 CATACTTCTCCGGACCATACGGGAGGACGTCGTGTCGTCGCAACGACTGGCACACCG 882

Db 883 GCCCTCTCTCGTGCTACCTCAAGAGCAACTACCAGTCCCACGGCATCTACAGGACGCAA 942

QY 665 AGTCGGTGCTGGCTATCCACAACATCGCCTTCCAGGCCGCATGTGGAGGAGGCTTCA 724

Best Local Similarity 57.2%; Pred. No. 1.4e-12;
Matches 210; Conservative 0; Mismatches 144; Indels 13; Gaps 2;

```
QY 172 GCGTGGACATCGTATGGTTGCTGCTGAGTGCCTTGGTCAACAGCGGCGCTG 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 GGGATGAACATAGTCTTTGTGGGCTGANATGGCTCCCTGGAGTAAACCGGAGACTC 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 GCGATGATGACTGGTGGCTGCCATTCAGCTGGTCAAGCGCGGCGCATGATGACC 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GGTGATGTTCTGGAGGACTGCCCGGCCATGGCTGCAATGGACAGAGTATGACT 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 ATTGCCCCCTGCTAGCACAGTACGCTGACGCTGGGACACTGGTGGTGGACATC 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 ATAGCTCACGCTATGATCAGTACAAAGATGGTGGGATACAAATGCTCTGGCTGAGTTA 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 ATGGCGGAGAGG-----TCCGCTACTTCCACTCCATCAAGAGGGCGTGCAC 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 AAAGTTGAAATGAAATTAAGANTCCGCTTCTTCCACTGTATAAAGAAAGANTTGAC 549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 CGCGTGTGGATGACACCCCTGTTCTGTCAGAGTCTGGGCAAGACCGGCTCCAAG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 AGGTTTTTCATGATCATCTTTGTTCTTGANAAGGTGTGGGAAAACCTGGAAGAATG 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 CTGACGCCCCCG-CTCGGCGCTGACTACTGTGACACACACAGCGCTTCGCCCTGTT 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 ATATATGTCCTCTGTCNAGGAACNGATTATGAANACACACAGCTAAGATTAACTTTTG 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 519 CTGCAAG 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 TTGCCAG 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8

US-10-145-415-98
; Sequence 98, Application US/10145415
; Publication No. US20030068788A1

GENERAL INFORMATION:

; APPLICANT: Buckel, Thomas Gunter
; APPLICANT: Hammer, Philip Eugene
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ligon, James Madison
; APPLICANT: Molnar, Istvan
; APPLICANT: Pachlatko, Johannes Paul
; APPLICANT: Zirkle, Ross Eric
; TITLE OF INVENTION: Methods and Compositions for Making Emamectin
; FILE REFERENCE: SYN-117 109846.312
; CURRENT APPLICATION NUMBER: US/10/145,415
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 60/291,149
; PRIOR FILING DATE: 2001-05-16

; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98

; LENGTH: 1266

; TYPE: DNA

; ORGANISM: Streptomyces

US-10-145-415-98

Query Match

Best Local Similarity 4.1%; Score 86.6; DB 9; Length 1266;

Matches 405; Conservative 0; Mismatches 509; Indels 3; Gaps 1;

```
QY 1204 GCCCTGGCCACAGATCTGGCCACCCCAAGTGTGAGTGGCATCTCTGGGTACCGCA-- 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GCGCAGCGCGCTCGGCTGCACCCCGAGGAGGAGCGCTCGGCTCGGCGACGCGACC 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1262 -AGGCGCGCTAGGAGTGTGAACGCCATCGCACCAAGTACAAGGCGCGCGCAAG 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 CTGATCGCTACACAAAGTGTGTGGCCACCGCGCGGAGAACCGCGGCTGGACATC 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1321 GCGGTGGTAAAGTTCTGCGCGCCCTGGCGACATGCTCAACGCGGCGCGGCGACTCATG 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 CCGCGCACGGGCTGGCGGCGGTGACACACCTGGCGCGCTCGGCCACGCGCGGCTG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9

US-09-893-519A-87

; Sequence 87, Application US/09893519A
; Publication No. US20030027243A1

GENERAL INFORMATION:

; APPLICANT: ANADYS PHARMACEUTICALS, INC.

; APPLICANT: THOMPSON, Craig

; APPLICANT: MOORE, Jeffrey

; APPLICANT: BUURMAN, Ed T.

; APPLICANT: BRADLEY, John

; APPLICANT: DESILVA, Thamara

; APPLICANT: HARRIS, Sandra

; APPLICANT: KOMARNITSKY, Svetlana

; APPLICANT: MENDILLO, Marc

; APPLICANT: MCCOY, Daniel

; APPLICANT: SANDERSON, Karen

; APPLICANT: HAQ, Tariq

; APPLICANT: ZHU, Shuhao

```
QY 1381 CTGTGTCCTCTCGGCTTCGAGCCCTTGGCGCTTGATCCAGCTGCAGCCATGCACTACGCT 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 CCGCGCTCTCTGGCCCTCCCTCGCGCGGACACAGCCCATCTGTTGATCGCCGAGCGCG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1441 ACCGTGCGCGTGTAGCTTCCACCGCGCGCTTGGTTCGACACCGTCAAGAGGCGGTCAAC 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 TGGATCGGCTGTGAGTGTCCCGCGCGCTTCTAGCGCGCGGAGTACCGCTGCTC 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1501 GCTTCCACATGGGCGCTGAACCCCGACAGCTTGGAGAGGCTGACGCCGACGCCCTG 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 GAGCGCGCCCGACCGCTGCACGGCATCTTGGGCGCCCGAACTCGCGGCTGTGTTCAAC 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1561 GCGCGCACCGTGCCTGCGCGTCCAGAGAGTGTTCGCGCGCGCGCTTCCACTTCGCGCGCTG 1620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 GATCTGCACCGGACACGGCTCGCTTCCACTTCGCGCGCGCTTCCACCGAGATCGCTC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1621 GCCAACTGATCAGCCAGACCTGCTTGGTTCGACCGCGCGCGCGCGCGCGCGCGCGCTG 1680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 GAGAGGCGCGCATGCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1681 CTGAGGAGGTGCTGTACGGCAAGGCGCGCTGGCCACCGCGCGCGCGCGCGCGCGCGCG 1740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 GATGTGCTCGCGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1741 GTGCGCGTTCGCGAGAGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 GATCTCGCGACCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1801 CTGAAGCGCTGCTCGCGCTTCCGTTGGAGGCAACGGCGCGCGCGCGCGCGCGCGCGCG 1860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 841 GACCGCTACATCTACGCGCGCGGTGACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1861 ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 901 CGGCTGCGGCTCGACACTTGGGCCAACGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCG 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1921 GCCACCGCGAGGTGACCACTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 961 ATGCTCGGCGAGGATCAGCTACGACCGCATCCGCTACTTCTCCGACGAGTACGAC 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1981 GCTGGCTCAAGCTGCGCGTGGCGCTCCACCTCGACCTCGGAGAACGCGCGCTGC 2040
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1021 GTCGGATGGATGACTCCCGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2041 TCCACGCGCACGCAACGCTGCGCTGCGCTTCCAGACCTCGGCTGCCAAGCGCGCTGTC 2100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1081 GCGGAGCTGCGCAAGCGGAGTTCATCGCTTCTGGCTGGCGCGCGCGCGCGCGCGCTC 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2101 TCCGCGCGCGCGCGCA 2117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1141 GCGGCGCATGAGGTCAA 1157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

APPLICANT: LONG, Pan
APPLICANT: DAVIDOV, Eugene
TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
FILE REFERENCE: 0342/IG548-US2
CURRENT APPLICATION NUMBER: US/09/893,519A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 87
LENGTH: 2307
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Human Genbank/Y11354.1
DATABASE ENTRY DATE: 1997-06-25
RELEVANT RESIDUES: (1)..(2307)
US-09-893-519A-87

Query Match 4.0%; Score 84.2; DB 9; Length 2307;
Best Local Similarity 45.7%; Pred. No. 2.6e-10;
Matches 450; Conservative 0; Mismatches 523; Indels 12; Gaps 4;

```
QY 1102 GCCCTGAGGCGGAGCTGGGCTGCTGTGGACATCATCTTGGCGGCTGCCCAAGATCTG 1161
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
70 GACCTGTGGGCTGCTGTGGAGTTCGAGGCTGGCGGCGAGCGCGGCCACCAACACCTC 129
QY 1162 GCGCGCTGGAGGAGCAAGAGGTGTGGACATCATCTTGGCGGCTGCCCAAGATCTG 1221
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
130 GCGCGCGGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 189
QY 1222 GCCACCCCAAGGTGCAGATGCCATCTCTGGGTACCGGCAAGGCGCGCTACAGAGTG 1281
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
190 AGCGGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 249
QY 1282 GTGAAGCCATCGGACCAAGTACAAAGGCGCGCGCAAGGCGGTGTGAAGTTCGCGC 1341
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
250 G-GAGCGCGCGGAGCGCGCGCGCGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCA 308
QY 1342 CCGCTGGCGCATCTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1401
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
309 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
QY 1402 CCCTGGCGCTGATCCAGCTGCACGCCATGACGTACGTACGCTGCGGTGGTACGCTCC 1461
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
368 CCGGGAAG-CTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
QY 1462 ACCGCGCGCTGTGCGACACCGTCAAGAGGCGGTCAAGGCTTCCACATGGGCGCGCTG 1521
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
427 GCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 1522 AACCCCGCAAGCTGGAGAGGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1581
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
487 CCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY 1582 AGCGAGGTGTTTGGGCGCGCGCTACCCCGAGATGTGGGCACTGATCAGCGCAGGAC 1641
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
547 GCGCGTGGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
QY 1642 CTGTCTGTGTCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1701
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
607 CTGAATCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 666
QY 1702 AAGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1761
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
667 CTGCTGCGCGTGCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
QY 1762 CCGCGCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1821
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
718 CCGTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
```

```
QY 1822 GTGAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1881
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
778 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
QY 1882 TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1941
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
838 CGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 1942 TACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2001
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
898 GCCAAGGTTATGCCAAGATCAGATTAAGCCAGAGCGGCGCGCGCGCGCGCGCGCG 957
QY 2002 GAGGCGCTCACCGCTCGACCTCGGAGAGCGCGCTGCCCAAGCGCGCGCGCGCG 2061
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
958 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017
QY 2062 GCCTCGGCGCTCCAAGACCTCGGCTG 2086
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
1018 GCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1042
```

RESULT 10

US-09-974-300-757
; Sequence 757, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 757
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-757

Query Match 3.9%; Score 83.2; DB 10; Length 1427;
Best Local Similarity 47.7%; Pred. No. 4.4e-10;
Matches 309; Conservative 0; Mismatches 333; Indels 6; Gaps 2;

```
QY 842 AGACCTACAAGAAGATCAACTGGCTGAAGGGTGGGCTATTCGCGCGCGCAAGCTGGTGA 901
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
578 AGTTTATGGGAACATTAATATATGAAGGGCGCATATCCCGCTGATCAGGTGACAA 637
QY 902 CTGTGTCGCGCAACTACGCGGACCGAGATC---GTCGCGGATGCGCGCGCGGTGTGAGC 958
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
638 CGGTACGTCCGACGTACCGCGATGAATCTTACTCTCTACTACGCGAGCGGCTTGAAG 697
QY 959 TGGACACCGTCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1018
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
698 GGTGCTCTGACAGAAAGGATGCACTGACAGGCAATTTAAACGGGATGACGATGTC 757
QY 1019 AGTGAAGCCCCAAGACCGCAAGTTCTGTCTGCGCGCTTACGACAGCAAGCGCTTACG 1078
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
758 TGTAGACCGCTTAAACGATCCGACATCGATTATCATACG---ATCCGCTCACCAGG 814
QY 1079 CCGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1138
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
815 ACGAAAAACGAAAAATAAAGCGGTATCCAGAGACATTCGGTCTGCGCGGTGAATGAG 874
QY 1139 CCGCGCGCGCTGTGCGCTTCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1198
Db ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
875 ACATTCCGCTCATCAGCATGTTGGCAAGGCTGACGAAGCAAAAAGGCTTCGATTGATCA 934
```


Db 11421 TTCCGAGGTCTATGACCGGAACCGGGCAACCAACCGCCCTACCGGACCACTCCGCG 11362
QY 1208 TGCCCAAGATCTTGGCCACCCCAAGTGTGAGATCGCCATCTCTGGGTACCGGAAGCGC 1267
Db 11361 ACCTCCCGCGCTCTCTGTCGCGACCAACGACCGCCACGCGCTCAACAACCAACGATACG 11302
QY 1268 CTTACGAGAGCTGGTGAACGCTATCGGACCAAGTACAAGGCGCGCCCAAGGCGTGG 1327
Db 11301 TGATCTGAGATCGAGGAGCCACCAACCGGATCCACCGGACCTCTGTATGAGGTCC 11242
QY 1328 TCAAGTCTTGGCGCCCTTGGCGACATGTCTACCGCGCGCGGCTTCAATGCTGTGTC 1387
Db 11241 TGAAGCCGAAGCGTGCACACCGCGCTTACTTCTGCGGGCTGCCACGAGTGGAGC 11182
QY 1388 CTTGCGGCTTCGAGCGCTCGGGGCTGATCCAGCTGACGCCATGCACTACGTTACCGTGC 1447
Db 11181 CGTACCGCGGCGACCGCGACCGCGCTGCGCGCTGCGCGACACCGAAGCTTGGACGCTCGC 11122
QY 1448 CGGTGTAGCTCTCACCGCGGCTGTGTGACACCTGACCGCGCGCGGCTTCAACGCGCTTC 1507
Db 11121 TGTCCCTGGACCGGACCGGCTGCGGACGAGCTGACCGCGCGGCTGCGGACCTGC 11062
QY 1508 ACATGGCGCCCTTGAACCCCGAAGCTGGACGAGCTGACCGCGCGGCTGCGGCGCA 1567
Db 11061 TGGGTCTTGGCGGACCGCGCGGCGAAGTACGCGCGCGGCTGCGGCGCGCGC 11002
QY 1568 CGGTGGCGCTGCCACGAGGTGTTTGGGGGCGCGCTTACCGCGGAGTGTGGCGCAACT 1627
Db 11001 CCGCGCTCGGCGCGCGC-----CAGACATCCACGCGCGCAAGTGGACGCTCCGA 10952
QY 1628 GCATCAGCAGGACCTGCTCTGTGTCAAGCGCGCGCGCGAGTGGAGGAGGAGTCAAGGTGCCG 1687
Db 10951 TGAACCGCGCGCGCTTTCGCGACCGCGCGCGCGAAGCTGCGCGCGCGCGCGCG 10892
QY 1688 AGGTGTGTAGCGCAAGGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1747
Db 10891 ATCTGGGGCGCGGTCCACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10832
QY 1748 TTGCGGAGAGATCCCGCGGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1807
Db 10831 CGACAGCGCGGAGACCG 10772
QY 1808 CGGTGTCCCGCTCTGTGAGGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1867
Db 10771 GCGCGTTCACCGCGTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10712
QY 1868 CGGCATGGGCGGTGGCG 1927
Db 10711 GCGCGCTCTCGCGCGCTTTCGCGGCGAGTCCCGGAGCTCCCGCGCGCGCGCGCGCG 10652
QY 1928 CCAAGGTACCACTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1987
Db 10651 CGCGCGGCGGTACTGGAAGAACCGCTCTCCCGTACAGGCTCGCGCGCGCGCGCGCG 10592
QY 1988 TCAAGCTGCGCGGTGAGGCTTCCACACCTCGACCTCGGAGAACCGCGCGCGCGCGCG 2047
Db 10591 CGCGCGTCCCGGAAGCGCGCTTCCCGTACAGGCTCGCGCGCGCGCGCGCGCGCGCG 10532
QY 2048 GCA 2050
Db 10531 GCA 10529

RESULT 13

US-09-860-846-3/c
; Sequence 3, Application US/09860846
; Patent No. US2002016472A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Liu, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/860.846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-3

Query Match 3.8%; Score 81; DB 9; Length 13613;
Best Local Similarity 42.7%; Pred. No. 1.3e-09;
Matches 488; Conservative 0; Mismatches 645; Indels 10; Gaps 1;

QY 908 CGCCCAACTACGCGACCGAGATGCTGCCGATGCCGCGGCGGTGTGAGCTGGACACCG 967
Db 12348 CGGTGACG 12349
QY 968 TCATCCGCGCGCAAGGGCATTTGAGGCGATTGTGACGCGCATGGACATTTGAGAGTGGAC 1027
Db 12288 CCAAGGCGGTCAACGCGCTTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12229
QY 1028 CCAAGACCGCAAAAGTTCTGTCTGCGCGCTTACGACCAAGACAGCGTCTAGCGCGCAAG 1087
Db 12228 CCGGATCCGCGCGCTTCCACAACTTCGCGCTTCGACCTGCGCGCGCGCGCGCGCG 12169
QY 1088 CCGCGCGCAAGGAGCGCTCGAGCGCGAGCTGGCGCTTGTGGAGCCCAACCGCGCGCG 1147
Db 12168 GGACCAACGCAAGATGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12109
QY 1148 TGTTCGCTTCATCGCGCGCGCTTGAGGAGCAGAGGTGTGGACATCATCTCTGCGCGCG 1207
Db 12108 TTCCCGAGGTCTCGACCGCGCAACCGCGCGCAACCGCGCGCGCGCGCGCGCGCGCG 12049
QY 1208 TGCCCAAGATCTCTGGCGCACCCCAAGGTGAGATCGCGATCTCTGGGTACCGCAAGCGCG 1267
Db 12048 ACCTCCCGCGCTCTCTGTCGCGCGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 11989
QY 1268 CTTACGAGAGCTGGTGAAGCGCATCGGACCAAGTACAAGGCGCGCGCGCGCGCGCGTG 1327
Db 11988 TGATCTGAGATCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11929
QY 1328 TCAAGTCTCGCGCGCGCTGGCGCACATGCTCAACCGCGCGCGCGCGCGCGCGCGCGCG 1387
Db 11928 TGAAGCGCGAAGCGGTGCACACCGCGCGCTTCTTCTGCGCGCGCTGCCAGAGCTGGAGC 11869
QY 1388 CTTGCGGCTTCGAGCGCTGGCGCTGATCCAGCTGACGCGCATGCACTACGTTACCGTGC 1447
Db 11868 CGTACCGCGGCG 11809
QY 1448 CCGTGTGAGCTCTCACCGCGCGCTGTGCACACCGTCAAGGAGGCGGTACCGGCTTCC 1507
Db 11808 TGTCCCTCG 11749
QY 1508 ACATGGCGCGCTTGAACCCCGCAAGCTTGAAGGCTGAGCGCGCGCGCGCGCGCGCGCGCG 1567
Db 11748 TGGCTCTCTGCG 11689
QY 1568 CCGTGGCGCGTGGCGAGGTTTGGCGGCGCGCGCTTACCGCGAGATGGTGGCGCAACT 1627
Db 11688 CCGCGCTCG 11639
QY 1628 GCATCAGCGAGACCTGCTCTGTTGCGGCGCGCGCGCTTACCGCGAGATGGTGGCGAG 1687
Db 11638 TGAACCG 11579
QY 1688 AGGTGGTGTACGCGAAGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1747
Db 11578 ATCTGGGCGCGGTTCCACGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11519

1748 TTGCCGAGAGATCCCGGCGACCTGCGCGCGCTGCTTCTAGCGCCCAACACCCCTGAAGC 1807
11518 OCAGAGGCGGAGAACGACCGCCCGCACCTCGTCCGGCTCGCGTGTAGGCAACA 11459
1808 CCGTGTCCGCTCGTGTGAGGAGAACGCGCGCGCGCCCAAGTGTGGGACACCGCCG 1867
11458 GCGCTTACACCCGCTGAGAGAGCGCCCGCACGACCTGGGGGTGACCGGAGCGCTTCC 11399
1868 CCGCATGGCGGCTGGCGCGACACCCCTCGGGCCCTCGCGCGCGCGCGCACCC 1927
11398 GCGGCTCTCGCTGTTCGGGAGGTCTCCGGAGTCCCGGACCGCGCTCGAGCGGCC 11339
1928 CCAAGTGACACCTTACAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCG 1987
11338 CCGCGGCGGCTACTTGAAGAACAACCTGCTCCCGCTCGAACAGCGCGCGCTTCGAGC 11279
1988 TCAAGTGGCGGCTGAGCGCTTCCACCACTCGACCTCGAGAGCGCGCTGCTTCAAGC 2047
11278 GCGGCTCGCGAGGAGCGCGCTTTCGGTACAGGCTCGGCTCTACCGCGCGCGCGCT 11219
2048 GCA 2050
11218 GCA 11216

RESULT 14
US-09-836-821-3/c
; Sequence 3, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-836-821-3

Query Match 3.8%; Score 81; DB 9; Length 13613;
Best Local Similarity 42.7%; Pred. No. 1.3e-09;
Matches 488; Conservative 0; Mismatches 645; Indels 10; Gaps 1;

908 GCGCCCACTACGCGACCGAGATCGCTGCGGATGCGCGCGCGGTGTGGAGTGGACACG 967
12348 CGGTGAGGCG 12289
968 TCATCCGCGCGCAAGGCGCATTTGAGGCGCATTTGAGGCGCATTTGAGGAGTGAAC 1027
12288 CCAAGGCGGTCAACGCTTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12229
1028 CCAAGACCGAGATTTCTGCTGCGCGCTAGACGAGAACAGAGTCTACGCGCGCAAG 1087
12228 CCGGATCG 12169
1088 CCGCGCGCAAGGAGCGCGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1147
12168 GAGCAACCGCAAGATGACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12109
1148 TGTTCGCTTCATCG 1207
12108 TTCCGAGGTCTATCGAGCGAAACCGCGCGCAACCGCGCGCTACCGGAGCACCTCGCG 12049
1208 TGCCCCAGATCTTGGCCACCCCGCAAGTGCAGATCGCCATCTGGGTACCGCGCAAGCG 1267

12048 ACCTCCCGGCGCTCTCGTTCGCGCGACACGACCGCGCGCGCTCAACACCGACGAGTACG 11989
1268 CCTACGAGAAGCTGTGTAAGCGCATCGGCAACCAAGTACAAGAGCGCGCGCAAGGCGGTGG 1327
11988 TGATCTCGATATCGACGAGGCGCACCGGATCAACCGCGACCTCTGTCATGAGGTTC 11929
1328 TCAAGTTCTCGCGCGCGCTTGGCGACATGCTCACCGCGCGCGCGCGCGCGCGCGCG 1387
11928 TGAAGCGGAGGCGTGCACACCGCGCGCTACTTTCGCGCGCGCTGCCACGAGCTGGAG 11869
1388 CCGTCCGCTTCGAGCGCTGCGCGCTGATCCACCTGCACCGCATGCACCGCATGCATGCTG 1447
11868 CGTACCGGCGGAGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11809
1448 CCGTGTAGCTTCCACCGCGCGCGCTGCTGCACACCGCTCAAGGAGGCGCGCTCACCGGTTC 1507
11808 TGTCCTCGCGACCG 11749
1508 ACATGGGCGCGCTGAACCG 1567
11748 TGCGTCTCTGCG 11689
1568 CCGTGGCGGCTGCGAGCGAGGTGTTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1627
11688 CCGCGCTCG 11639
1628 GCATCAGCGAGGACCTGCTCTGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1687
11638 TGACCG 11579
1688 AGTGTGTACGCGAGGCG 1747
11578 ATCTGGGCG 11519
1748 TTCCGAGAGATCCCG 1807
11518 CCGAGGAGCGCGGAGCG 11459
1808 CCGTGTCCGCTTCGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1867
11458 GCGCGTTCACCGCGCTGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11399
1868 CCGCATGCGCGGTGGCG 1927
11398 GCGCGCTCTCTCGCGCTTTCG 11339
1928 CCAAGGTGACCACTTACAAGCG 1987
11338 CCGCGGCGCGGTACTTGAAGAACACCGCTGCTCCGCTCGAACAGCGCGCGCGCTTCGAGC 11279
1988 TCAAGTGGCGGTGAGGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2047
11278 GCGCGCTCGCGAGGAGCGCGCTTCCCGCTACAGCGCTCGCGCGCGCGCGCGCGCGCGCG 11219
2048 GCA 2050
11218 GCA 11216

RESULT 15
US-09-861-289-3/c
; Sequence 3, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18

This Page Blank (uspto)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, .rnpm and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, .rapm and .rapn

The Pending database search results should not be left in the case because they contain data that is confidential.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:50:00 ; Search time 3310.08 Seconds
(without alignments)
16133.327 Million cell updates/sec

Title: US-09-980-771A-2
Perfect score: 2124
Sequence: 1 atgtgtgttcctaccag.....cgccaccgcaagtcgcgc 2124

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues 49582208
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	2124	100.0	2124	37 US-09-980-771A-2	Sequence 2, Appli
2	2124	100.0	3117	37 US-09-980-771A-1	Sequence 1, Appli
3	1953	91.9	1953	37 US-09-980-771A-4	Sequence 4, Appli
4	1593	75.0	1593	37 US-09-980-771A-8	Sequence 8, Appli
5	1314	61.9	1314	37 US-09-980-771A-6	Sequence 6, Appli
6	533.8	25.1	2058	3 US-07-771-351-4	Sequence 4, Appli
7	527.4	24.8	2201	3 US-07-771-351-5	Sequence 5, Appli
8	523.8	24.7	2279	3 US-07-771-351-3	Sequence 3, Appli
9	478	22.5	1818	29 US-09-731-166-3	Sequence 1, Appli
10	478	22.5	2267	7 US-08-300-726-1	Sequence 1, Appli
11	478	22.5	2267	7 US-08-300-726A-1	Sequence 25, Appli
12	478	22.5	2267	36 US-09-961-077-25	Sequence 46, Appli
13	472.8	22.3	1915	18 US-09-402-254-46	Sequence 385561,
14	449.2	21.1	2566	25 US-09-654-617-385561	Sequence 385561,
15	449.2	21.1	2566	27 US-09-684-016-385561	Sequence 6, Appli
16	447.6	21.1	2542	24 US-07-771-351-6	Sequence 1, Appli
17	323.2	15.2	1160	3 US-07-771-351-1	Sequence 1, Appli
18	314.8	14.8	2399	40 US-10-138-075-1	Sequence 9995, Ap
19	314.8	14.8	2399	72 US-60-288-315-1	Sequence 258750,
20	305.2	14.4	2098	42 US-10-219-999-9995	
21	305.2	14.4	2284	25 US-09-654-617-258750	

QY 1321 GCGTGGTCAAGTCTCGCGCGCCCTCGCGGCACATGCTACCGCGCGCGCCGACATTCATG 1380
Db 1321 GCGTGGTCAAGTCTCGCGCGCCCTCGCGGCACATGCTACCGCGCGCGCCGACATTCATG 1380
QY 1381 CTGGTGGCCCTCGCGCTTCGAGCCCTCGCGCCCTGATCCAGCTGACGCCATGACATCGGT 1440
Db 1381 CTGGTGGCCCTCGCGCTTCGAGCCCTCGCGCCCTGATCCAGCTGACGCCATGACATCGGT 1440
QY 1441 ACCGTGCCGTGGTGGTCCAGCCCTCCAGCGCGCCCTGGTGCACACCTGAAGGAGCGCGTCAAC 1500
Db 1441 ACCGTGCCGTGGTGGTCCAGCCCTCCAGCGCGCCCTGGTGCACACCTGAAGGAGCGCGTCAAC 1500
QY 1501 GCTTCCACATGCGCGCCCTGTAACCCCGACAGCTGGAGCGCTGACGCCGACGCCCTG 1560
Db 1501 GCTTCCACATGCGCGCCCTGTAACCCCGACAGCTGGAGCGCTGACGCCGACGCCCTG 1560
QY 1561 GCGGCCACGTGCGCGCTGCGAGCGAGGTGTTGGGGGCGCGCTACCCCGAGATGG 1620
Db 1561 GCGGCCACGTGCGCGCTGCGAGCGAGGTGTTGGGGGCGCGCTACCCCGAGATGG 1620
QY 1621 GCAACTGCATCAGCCAGGACCTGCTCTGCTCCAGCCCGCCAGAGTGGAGGCGCTG 1680
Db 1621 GCAACTGCATCAGCCAGGACCTGCTCTGCTCCAGCCCGCCAGAGTGGAGGCGCTG 1680
QY 1681 CTGGAGGAGTGGTGTACGGGAAGGGCGCGCTGGCCACCGCCAGAGGAGGATCAAG 1740
Db 1681 CTGGAGGAGTGGTGTACGGGAAGGGCGCGCTGGCCACCGCCAGAGGAGGATCAAG 1740
QY 1741 GTGGCCGTTCGCGAAGATCCCGCGGACCTGCGCGCGCTGCTACGCCGCCAACACC 1800
Db 1741 GTGGCCGTTCGCGAAGATCCCGCGGACCTGCGCGCGCTGCTACGCCGCCAACACC 1800
QY 1801 CTGAAGCCCGTGTCCGCTCCGTTGGAGGCAACGGCGCGCGCGCCGCGCCAAAGTTCGCAAC 1860
Db 1801 CTGAAGCCCGTGTCCGCTCCGTTGGAGGCAACGGCGCGCGCGCCGCGCCAAAGTTCGCAAC 1860
QY 1861 ACCGCCCGCATGGCGCGTGGCGCGGACCAACCCCTCGGGCCCTCGCGCGCGCG 1920
Db 1861 ACCGCCCGCATGGCGCGTGGCGCGGACCAACCCCTCGGGCCCTCGCGCGCGCG 1920
QY 1921 GCCACCCCAAGTGACACCTACAAAGCCCGCTCGCGCGCACCGCCCAAGCCCAAGACC 1980
Db 1921 GCCACCCCAAGTGACACCTACAAAGCCCGCTCGCGCGCACCGCCCAAGCCCAAGACC 1980
QY 1981 GCTGGCCTCAAGTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGCGCTGCC 2040
Db 1981 GCTGGCCTCAAGTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGCGCTGCC 2040
QY 2041 TCCAGGCAACGGCAACGGTGCCTCGGCTTCCAGACCTCGGCTGCCAAGCCCTGGTG 2100
Db 2041 TCCAGGCAACGGCAACGGTGCCTCGGCTTCCAGACCTCGGCTGCCAAGCCCTGGTG 2100
QY 2101 TCCGCGCCACCGCAAGTCCGCG 2124
Db 2101 TCCGCGCCACCGCAAGTCCGCG 2124

RESULT 2
US-09-980-771a-1
; Sequence 1, Application US/09980771A
; GENERAL INFORMATION:
; APPLICANT: CNRS
; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTEREST,
; FILE OF INVENTION: METHOD OF OBTAINING THEM AND THEIR USES
; FILE REFERENCE: WO 99 AB CNR AMYL
; CURRENT APPLICATION NUMBER: US/09/980, 771A
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: FR 99/06494
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

; LENGTH: 3117
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-980-771a-1
Query Match 100.0%; Score 2124; DB 37; Length 3117;
Best Local Similarity 100.0%; Pred. No. 1.6e-292;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTGTGGCTCTACACGCCGCCAGCGCGCTCTATCGTTCATCAATGCGCG 60
Db 15 ATGGCTGTGGCTCTACACGCCGCCAGCGCGCTCTATCGTTCATCAATGCGCG 74
QY 61 TCGTTCGGTGTCAAGAGAACCGGAAACAGCTGCTGCTGAGTTCCTGCTGCTCCGCA 120
Db 75 TCGTTCGGTGTCAAGAGAACCGGAAACAGCTGCTGCTGAGTTCCTGCTGCTCCGCA 134
QY 121 CGCAAGTCCACCTCGCGCTCGGCTGTACTGGTGCCACTGGTGCCACTTCGCGCGTGGAC 180
Db 135 CGCAAGTCCACCTCGCGCTCGGCTGTACTGGTGCCACTGGTGCCACTTCGCGCGTGGAC 194
QY 181 ATCGTGTGCTGCTGCTGAGTTCGCGCTTGTGTCGAAGACGGCGCGCTGGCGGATGTG 240
Db 195 ATCGTGTGCTGCTGCTGAGTTCGCGCTTGTGTCGAAGACGGCGCGCTGGCGGATGTG 254
QY 241 ACTGTGGCTGCTGCTGAGTTCGCGCTGAGCGCGCGCTGCGGCTCATGACCATTCGCCCT 300
Db 255 ACTGTGGCTGCTGCTGAGTTCGCGCTGAGCGCGCGCTGCGGCTCATGACCATTCGCCCT 314
QY 301 CGCTAGCAGCAGTACCTGACCTGCGGACACCTCGGTGGTGCTGGACATCATGGCGGAG 360
Db 315 CGCTAGCAGCAGTACCTGACCTGCGGACACCTCGGTGGTGCTGGACATCATGGCGGAG 374
QY 361 AAGTCCGCTACCTCCACTCCATCAAGAGGGGTGCACCGCGCTGTGGTGTGATGACCCGCC 420
Db 375 AAGTCCGCTACCTCCACTCCATCAAGAGGGGTGCACCGCGCTGTGGTGTGATGACCCGCC 434
QY 421 TGTGCTGCTGGCAAGGTCTGGGGCAGACCGGCTTCAAGCTGACGGCGCGCGCTGCGGC 480
Db 435 TGTGCTGCTGGCAAGGTCTGGGGCAGACCGGCTTCAAGCTGACGGCGCGCGCTGCGGC 494
QY 481 GCTGACTACCTGACACACACACAGCGCTTCGCCCTGTCTGCAAGCGCGCTATGAGGCT 540
Db 495 GCTGACTACCTGACACACACACAGCGCTTCGCCCTGTCTGCAAGCGCGCTATGAGGCT 554
QY 541 GCGCGGTGCTGCTGCTGCGCGCGCGGAGTGGTCTTCCTGCGCAACAGCTGSCAC 600
Db 555 GCGCGGTGCTGCTGCTGCGCGCGCGGAGTGGTCTTCCTGCGCAACAGCTGSCAC 614
QY 601 TCGCGCTGCTGCTGCTGCTGCTGAGGAGAGTACAGCCCAAGGGCGAGTTCACCAAG 660
Db 615 TCGCGCTGCTGCTGCTGCTGCTGAGGAGAGTACAGCCCAAGGGCGAGTTCACCAAG 674
QY 661 GCCAAGTGGTGTGCTATCCACACATCCCTTCCAGGGCGCGCATGTGGGAGGAGGCT 720
Db 675 GCCAAGTGGTGTGCTATCCACACATCCCTTCCAGGGCGCGCATGTGGGAGGAGGCT 734
QY 721 TTCAAGGACACGAAGTGGCGCGCGCTTTTGAACAAGTGGCTTCCTGCGAGCGCTAT 780
Db 735 TTCAAGGACACGAAGTGGCGCGCGCTTTTGAACAAGTGGCTTCCTGCGAGCGCTAT 794
QY 781 GCCAAGTGTACTACTGAGGCCACCCCTGAGGAGAGACAGAGGCCCGCTGACGGGA 840
Db 795 GCCAAGTGTACTACTGAGGCCACCCCTGAGGAGAGACAGAGGCCCGCTGACGGGA 854
QY 841 AAGACCTACAAGAGATCAACTGCTGAAGGGTGGCATTCATCCCGCCGACAGCTGGTG 900
Db 855 AAGACCTACAAGAGATCAACTGCTGAAGGGTGGCATTCATCCCGCCGACAGCTGGTG 914
QY 901 ACTGTGTCGCCCAACTACGCGAGATCGCTGCGGATGCCCGCGCGGTGTGGAGCTG 960
Db 915 ACTGTGTCGCCCAACTACGCGAGATCGCTGCGGATGCCCGCGCGGTGTGGAGCTG 974

```
QY 961 GACACGTCATCGCGCCCAAGGGCATTTGAGGGCATTTGAAAGGCAATGACATTTGAGGAG 1020
Db 975 GACACGTCATCGCGCCCAAGGGCATTTGAGGGCATTTGAAAGGCAATGACATTTGAGGAG 1034
QY 1021 TGGAAACCCCAAGACCAAGATTCCTGCTGCGCCCTAGCACCAGACAGCGCTCTAGGCC 1080
Db 1035 TGGAAACCCCAAGACCAAGATTCCTGCTGCGCCCTAGCACCAGACAGCGCTCTAGGCC 1094
QY 1081 GGCAGGCGCGCCCAAGGAGGCGCTGCAAGGCGAGCTGGCGCTGCTGGAACCCCAAC 1140
Db 1095 GGCAGGCGCGCCCAAGGAGGCGCTGCAAGGCGAGCTGGCGCTGCTGGAACCCCAAC 1154
QY 1141 GCGCCCTGCTGCGCTTCAATGCGCGCCCTGAGGAGGAGCAAGAGGTGAGACATCATCTG 1200
Db 1155 GCGCCCTGCTGCGCTTCAATGCGCGCCCTGAGGAGGAGCAAGAGGTGAGACATCATCTG 1214
QY 1201 GCGCGCTGCTGCGCAAGATCTGCGCACCCCAAGGTGACATGCGCATCTGGGTACCGGC 1260
Db 1215 GCGCGCTGCTGCGCAAGATCTGCGCACCCCAAGGTGACATGCGCATCTGGGTACCGGC 1274
QY 1261 AAGGCGCGCTAGGAGAGCTGCTGAACGCGCATCGGCAACAGTACAAAGGCGCGGCAAG 1320
Db 1275 AAGGCGCGCTAGGAGAGCTGCTGAACGCGCATCGGCAACAGTACAAAGGCGCGGCAAG 1334
QY 1321 GCGGTGCTCAAGTCTCGCGCGCCCTGCGCGCACATGCTCAACGCGCGCGGCGGACTCATG 1380
Db 1335 GCGGTGCTCAAGTCTCGCGCGCCCTGCGCGCACATGCTCAACGCGCGCGGCGGACTCATG 1394
QY 1381 CTGGTGCCTGCTGCGCTTCAAGCGCTGCGCGCATGATCCAGCTGACACGCGCATGACTACGT 1440
Db 1395 CTGGTGCCTGCTGCGCTTCAAGCGCTGCGCGCATGATCCAGCTGACACGCGCATGACTACGT 1454
QY 1441 ACCGTGCGCGTGGTAGCTTCCACGCGCGCTGCTGGTGACACGCTCAAGGAGGCGGTCAAC 1500
Db 1455 ACCGTGCGCGTGGTAGCTTCCACGCGCGCTGCTGGTGACACGCTCAAGGAGGCGGTCAAC 1514
QY 1501 GCTTCCACATGCGCGCTGCAACCGCGAGAGCTGGAGAGGCTGACGCGCGCGGCTG 1560
Db 1515 GCTTCCACATGCGCGCTGCAACCGCGAGAGCTGGAGAGGCTGACGCGCGCGGCTG 1574
QY 1561 GCGCGCACGCTGCGCGTGCACGAGGCTGTTGGGCGCGCGCTTACCGCGAGATGCTG 1620
Db 1575 GCGCGCACGCTGCGCGTGCACGAGGCTGTTGGGCGCGCGCTTACCGCGAGATGCTG 1634
QY 1621 GCGCACTGATCAGCAGGAGCTGCTGGTCCAGCGCGCGCGCGCGCGCGCGCGCGCTG 1680
Db 1635 GCGCACTGATCAGCAGGAGCTGCTGGTCCAGCGCGCGCGCGCGCGCGCGCGCGCTG 1694
QY 1681 CTGGAGGAGGCTGCTACGCGCAAGGCGCGCTGGCGCACCGCGCAAGGAGGAGATCAAG 1740
Db 1695 CTGGAGGAGGCTGCTACGCGCAAGGCGCGCTGGCGCACCGCGCAAGGAGGAGATCAAG 1754
QY 1741 GTGCGCGTTCGCGAGAGATCCCGCGGACCTGCGCGCGCGCTGCTTACCGCGCGCGCG 1800
Db 1755 GTGCGCGTTCGCGAGAGATCCCGCGGACCTGCGCGCGCGCTGCTTACCGCGCGCGCG 1814
QY 1801 CTGAAGCGCGTGTGCGCTCCGTGGAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCTG 1860
Db 1815 CTGAAGCGCGTGTGCGCTCCGTGGAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCTG 1874
QY 1861 ACCGCGCGCGCTGCGCGCGCTGGCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1920
Db 1875 ACCGCGCGCGCTGCGCGCGCTGGCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1934
QY 1921 GCGACCCCAAGGTGACCACTTCAAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 1980
Db 1935 GCGACCCCAAGGTGACCACTTCAAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 1994
QY 1981 GCTGCGCTCAAGCTGGCGGCTGAGCGCTTCACCACTGACCTCGGAGAAAGCGCGCTGCC 2040
Db 1995 GCTGCGCTCAAGCTGGCGGCTGAGCGCTTCACCACTGACCTCGGAGAAAGCGCGCTGCC 2054
QY 2041 TCCAAAGCGCAAGGCGGCTGCTGCGCTTCCAAAGCTGCTGCGCGCGCGCGCGCGCTG 2100
```

```
Db 2055 TCCAAAGCGCAAGGCGGCTGCTGCGCTTCCAAAGCTGCTGCGCTGCGCAAGCGCGCTGTC 2114
QY 2101 TCCGCGCGCACCGCGCAAGTCCGCC 2124
Db 2115 TCCGCGCGCACCGCGCAAGTCCGCC 2138

RESULT 3
US-09-980-771A-4
; Sequence 4, Application US/09980771A
; GENERAL INFORMATION:
; APPLICANT: CNRS
; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTERE
; FILE REFERENCE: WO 99 AB CNR AMYL
; CURRENT APPLICATION NUMBER: US/09/980, 771A
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: FR 99/06494
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: fragment of the complete sequence of cDNA coding for the GBSSI
; OTHER INFORMATION: Chlamydomonas reinhardtii and coding for the mature GBSSI prot
; OTHER INFORMATION: n
; NAME/KEY: CDS
; LOCATION: (1)..(1953)
; OTHER INFORMATION:
US-09-980-771A-4
```

```
Query Match 91.9%; Score 1953; DB 37; Length 1953;
Best Local Similarity 100.0%; Pred. No. 3.3e-268;
Matches 1953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCGCTGCGACATCGTATGTTGCTGAGTGCCTGAGTGCCTTGGTCCCAAGAGCGGCGGCTG 231
Db 1 GCGCTGCGACATCGTATGTTGCTGCTGAGTGCCTTGGTCCCAAGAGCGGCGGCTG 60
QY 232 GCGGATGTACTGTTGGCTGCTGCTATGAGTGGTCAAGCGGCGGCGGCGGCTGATGACC 291
Db 61 GCGGATGTACTGTTGGCTGCTGCTATGAGTGGTCAAGCGGCGGCGGCGGCTGATGACC 120
QY 292 ATTGCCCTCGCTTACACACAGTACGCTGAGCGCTGGGACACCTCGGTTGCTGGGACATC 351
Db 121 ATTGCCCTCGCTTACACACAGTACGCTGAGCGCTGGGACACCTCGGTTGCTGGGACATC 180
QY 352 ATGGCGGAGAAGTCCGCTACTTCCACTCCTCAAGAGGCGGCTGACCGCGTGTGGATT 411
Db 181 ATGGCGGAGAAGTCCGCTACTTCCACTCCTCAAGAGGCGGCTGACCGCGTGTGGATT 240
QY 412 GACCAACCCCTGGTTCTGGCAAGGCTCTGGGCAAGCGGCTGCAAGCTGTACGCGCCC 471
Db 241 GACCAACCCCTGGTTCTGGCAAGGCTCTGGGCAAGCGGCTGCAAGCTGTACGCGCCC 300
QY 472 CGCTCGGCGCTGACTTACCTGAGCAACACGCGCTTCCGCTGCTTCTGCAAGGCGCT 531
Db 301 CGCTCGGCGCTGACTTACCTGAGCAACACGCGCTTCCGCTGCTTCTGCAAGGCGCT 360
QY 532 ATTGAGGCTCGCGGCTGCTGCTTCCGCGCGCGGCGGAGGACTGCTTCTGCGGCAAC 591
Db 361 ATTGAGGCTCGCGGCTGCTGCTTCCGCGCGCGGCGGAGGACTGCTTCTGCGGCAAC 420
QY 592 GACTGGCACTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
Db 421 GACTGGCACTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 652 TTCACCAAGCGCAAGTTCGCTGCTGCTATCCACACATCGCTTCCAGGCGCGGCTGCTG 711
```

```
Db 481 TTCACCAAGCCCAAGTCGGTGGCTATCCAAACATCGCTTCCAGGGCCGATGG 540
QY 712 GAGGAGGCTTTCAGGACAGAAAGCTGCCCGAGCGCCCTTTGACAAAGCTGGCTTCTCG 771
Db 541 GAGGAGGCTTTCAGGACAGAAAGCTGCCCGAGCGCCCTTTGACAAAGCTGGCTTCTCG 600
QY 772 GAGGCTATGCGAAGGTTTACACTGAGGCCACCCCAATGAGGAGGACGAGAGCCCGCG 831
Db 601 GAGGCTATGCGAAGGTTTACACTGAGGCCACCCCAATGAGGAGGACGAGAGCCCGCG 660
QY 832 CTGACGGGAAGACCTACAAGAAGATCAACTGGCTGAAGGTGCAATTATCGCGCGGAC 891
Db 661 CTGACGGGAAGACCTACAAGAAGATCAACTGGCTGAAGGTGCAATTATCGCGCGGAC 720
QY 892 AAGCTGTGACTGTGTGCGCCCAACTACGCGACCGAGATCGTCCGATGCGCGCGGCT 951
Db 721 AAGCTGTGACTGTGTGCGCCCAACTACGCGACCGAGATCGTCCGATGCGCGCGGCT 780
QY 952 GTGGAGCTGACACCGCTATCCGGCCCAAGGGCAATTGAGGGCAATTGTAAGGGATGGAC 1011
Db 781 GTGGAGCTGACACCGCTATCCGCGCCCAAGGGCAATTGAGGGCAATTGTAAGGGATGGAC 840
QY 1012 ATTGAGGAGTGGAAACCCCAAGACCGACAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC 1071
Db 841 ATTGAGGAGTGGAAACCCCAAGACCGACAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC 900
QY 1072 GTCTACGCGGCAAGGCGCGGCCCAAGGAGGSCCTGCAGCGCGCTGGGCTGCCCTGTG 1131
Db 901 GTCTACGCGGCAAGGCGCGGCCCAAGGAGGSCCTGCAGCGCGCTGGGCTGCCCTGTG 960
QY 1132 GACCCCAAGGCGCGGCCCTGTGCGCTTTCATGCGCGCCCTGAGGAGGAGAGGGTGTGGAC 1191
Db 961 GACCCCAAGGCGCGGCCCTGTGCGCTTTCATGCGCGCCCTGAGGAGGAGAGGGTGTGGAC 1020
QY 1192 ATCATCTGCGCGCCCTGCCCAAGATCTGCGCCACCCCAAGTGCAGATCGCATCTG 1251
Db 1021 ATCATCTGCGCGCCCTGCCCAAGATCTGCGCCACCCCAAGTGCAGATCGCATCTG 1080
QY 1252 GGTACCGGCAAGGCGCGCCCTAGGAGAAGTGGTGAACGCCAATPCGCAACCAAGTACAGGGC 1311
Db 1081 GGTACCGGCAAGGCGCGCCCTAGGAGAAGTGGTGAACGCCAATPCGCAACCAAGTACAGGGC 1140
QY 1312 CGCGCCAAAGCGGTGCTCAAGTTCCTGCGCGCCCTGGCGACATGCTCACCGCGCGGCC 1371
Db 1141 CGCGCCAAAGCGGTGCTCAAGTTCCTGCGCGCCCTGGCGACATGCTCACCGCGCGGCC 1200
QY 1372 GACTTTCATGTGTGTCGCTCGCGCTTCGAGCGCCCTGCGGCTGATCCAGCTGCAGCGCCATG 1431
Db 1201 GACTTTCATGTGTGTCGCTCGCGCTTCGAGCGCCCTGCGGCTGATCCAGCTGCAGCGCCATG 1260
QY 1432 CACTACCGGTACCGTGGCCGTGGTAGCCTCCACCGGGCGGCTGTGTCGACACCGTCAAGGAG 1491
Db 1261 CACTACCGGTACCGTGGCCGTGGTAGCCTCCACCGGGCGGCTGTGTCGACACCGTCAAGGAG 1320
QY 1492 GGGCTACCGGCTTCCACATGGGCGCCCTGAACCCCGACAAAGCTGGAGGAGGCTGACGCC 1551
Db 1321 GGGCTACCGGCTTCCACATGGGCGCCCTGAACCCCGACAAAGCTGGAGGAGGCTGACGCC 1380
QY 1552 GACGCCCTGGCGCCACCGTGCAGCGCTGCGCAGAGGTGTTTGGGGGGCGGCTACGCC 1611
Db 1381 GACGCCCTGGCGCCACCGTGCAGCGCTGCGCAGAGGTGTTTGGGGGGCGGCTACGCC 1440
QY 1612 GAGATGTGGCCAACTGTCATCAGCCAGGACCTGTCTGGTCCAAAGCCCGCCCAAGATGG 1671
Db 1441 GAGATGTGGCCAACTGTCATCAGCCAGGACCTGTCTGGTCCAAAGCCCGCCCAAGATGG 1500
QY 1672 GAGGCGCTGTGGAGGAGTGGTACGGCAAGGCGCGCTGGCCACCGCCCAAGAGAG 1731
Db 1501 GAGGCGCTGTGGAGGAGTGGTACGGCAAGGCGCGCTGGCCACCGCCCAAGAGAG 1560
QY 1732 CAGATCAAGGTGCGCGTGGCGGAGAAGATCCCGGCGACCTGCGCGCGGCTGTCTACGCC 1791
Db 1561 GAGATCAAGGTGCGCGTGGCGGAGAAGATCCCGGCGACCTGCGCGCGGCTGTCTACGCC 1620
```

```
QY 1792 CCCAACACCCCTGAAGCCCGTGTCCGCTTCCGTGAGGGCAACGGCGCGCGCCGCCAAG 1851
Db 1621 CCCAACACCCCTGAAGCCCGTGTCCGCTTCCGTGAGGGCAACGGCGCGCGCCGCCAAG 1680
QY 1852 GTCCGGCACCCGCGCGCCCGCCATGGCGCGTGGCGGCGGAGACACCCCTCGGGCCCTCG 1911
Db 1681 GTCCGGCACCCGCGCGCCCGCCATGGCGCGTGGCGGCGGAGACACCCCTCGGGCCCTCG 1740
QY 1912 CCCCGCGCGCACCCCGCCCAAGGTGACCACTTACAAAGCCCGCTTCCGCGCGCCACCGCCAAAG 1971
Db 1741 CCCCGCGCGCACCCCGCCCAAGGTGACCACTTACAAAGCCCGCTTCCGCGCGCCACCGCCAAAG 1800
QY 1972 CCCAAGACCGCTGCGCTCAAGCTGGCGGCTGAGGCTTCCACCACTCGACCTCGGAGAAC 2031
Db 1801 CCCAAGACCGCTGCGCTCAAGCTGGCGGCTGAGGCTTCCACCACTCGACCTCGGAGAAC 1860
QY 2032 GCGCTGCTTCCAAAGCGCAAGCGCAAGCGTGGCTTGGCTTCCAAAGCTTGGCTGCCAAG 2091
Db 1861 GCGCTGCTTCCAAAGCGCAAGCGCAAGCGTGGCTTGGCTTCCAAAGCTTGGCTGCCAAG 1920
QY 2092 CCGCTGCTTCCGCGCGCACCGCGCAAGTCCGCC 2124
Db 1921 CCGCTGCTTCCGCGCGCACCGCGCAAGTCCGCC 1953

RESULT 4
US-09-980-771A-8
; Sequence 8, Application US/09980771A
; GENERAL INFORMATION:
; APPLICANT: CNRS
; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF INTER
; FILE REFERENCE: WO/99 AB CNR AMYL
; CURRENT APPLICATION NUMBER: US/09/380,771A
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: FR 99/06494
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: fragment of the complete cDNA coding for the GBSSI of Chlamy
; OTHER INFORMATION: as reinhardtii
; NAME/KEY: CDS
; LOCATION: (1)..(1593)
; OTHER INFORMATION:
US-09-980-771A-8

Query Match 75.0%; Score 1593; DB 37; Length 1593;
Best Local Similarity 100.0%; Pred. No. 4.5e-217;
Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 GCGCTGGACATCGGTGATGGTGTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGGCGGCTG 231
Db 1 GCGCTGGACATCGGTGATGGTGTGCTGCTGAGGTGCGCCCTTGGTCCAAAGACGGGCGGCTG 60
QY 232 GCGCATGTGACTGTGGGCTGCCTATTGAGCTGGTCAAGCGCGGCCACCGGTCATGACC 291
Db 61 GCGCATGTGACTGTGGGCTGCCTATTGAGCTGGTCAAGCGCGGCCACCGGTCATGACC 120
QY 292 ATTGCCCTCGCTACGACCACTACGCTGACCGCTGGGACACCTCGTGGTGGTGGACATC 351
Db 121 ATTGCCCTCGCTACGACCACTACGCTGACCGCTGGGACACCTCGTGGTGGTGGACATC 180
QY 352 ATGGCGGAGAGGTCGCTACTTCCATCAAGAGAGGCGGTGCACCGGCTGTGGATT 411
Db 181 ATGGCGGAGAGGTCGCTACTTCCATCCATCAAGAGAGGCGGTGCACCGGCTGTGGATT 240
QY 412 GACCACCCCTGGTCTCTGGCCAAAGTCTGCGGCAAGACCGGCTCCAAAGCTGTACGGCCCC 471
```

Db	241	GACCACCCCTGGTTCTCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAGCTGTACGGCCCC	300
QY	472	CGCTCGGGCGCTGACTACTCTGGACAAACCAAGCGCTTCGCCCTGTGTTCTGCAAGGCGCGT	531
Db	301	CGTCCGGCGCTGACTACTCTGGACAAACCAAGCGCTTCGCCCTGTGTTCTGCAAGGCGCGT	360
QY	532	ATTGAGGCTCGCGCGTCTGCCCTTCGGCCCGCGGAGGACATGCGTCTTCGTGGCCCAAC	591
Db	361	ATTGAGGCTCGCGCGTCTGCCCTTCGGCCCGCGGAGGACATGCGTCTTCGTGGCCCAAC	420
QY	592	GACTGGCACTCCGCCCTGTGTCGGTCTGAAAGGACGAGTACCAAGCCCAAGGGCGAG	651
Db	421	GACTGGCACTCCGCCCTGTGTCGGTCTGAAAGGACGAGTACCAAGCCCAAGGGCGAG	480
QY	652	TTACCAAGGCCAAGTCGGTCTGGCTATCCACAACATCGCTTCCAGGGCGCGCATGTGG	711
Db	481	TTACCAAGGCCAAGTCGGTCTGGCTATCCACAACATCGCTTCCAGGGCGCGCATGTGG	540
QY	712	GAGGAGGCTTTCAAGGACACAAAGCTGCCCCAGGCCGCTTTGACAAGCTGGCCTTCTCG	771
Db	541	GAGGAGGCTTTCAAGGACACAAAGCTGCCCCAGGCCGCTTTGACAAGCTGGCCTTCTCG	600
QY	772	GAGGCTATGCCAAGGTTTACATGAGGCCACCCCCATGAGGAGGACGAGAAGCCCCCG	831
Db	601	GAGGCTATGCCAAGGTTTACATGAGGCCACCCCCATGAGGAGGACGAGAAGCCCCCG	660
QY	832	CTGACGGGAAGACCTACAGAGACATCACTGCTGAAGGTTGGCATTTATCGCCGCGCAG	891
Db	661	CTGACGGGAAGACCTACAGAGACATCACTGCTGAAGGTTGGCATTTATCGCCGCGCAG	720
QY	892	AAGCTGCTGACTGTGTGCCCAACTACGCGACCGAGATCGCTGCGCATGCGCGCGCGT	951
Db	721	AAGCTGCTGACTGTGTGCCCAACTACGCGACCGAGATCGCTGCGCATGCGCGCGCGT	780
QY	952	GTGGAGCTGACACCGTATCCGCGCAAGGGCATTTGAGGCGATTGTAACGGCATGGAC	1011
Db	781	GTGGAGCTGACACCGTATCCGCGCAAGGGCATTTGAGGCGATTGTAACGGCATGGAC	840
QY	1012	ATTGAGAGTGGAAACCCCAAGCCAGCAAGTCTCTGCTCGGCCCTACGACCAAGACAGC	1071
Db	841	ATTGAGAGTGGAAACCCCAAGCCAGCAAGTCTCTGCTCGGCCCTACGACCAAGACAGC	900
QY	1072	GTCTAGCGCGGAAGGCGCGCGCAAGAGGCGCTTCGACGCGAGCTGGGCCCTGCGTGTG	1131
Db	901	GTCTAGCGCGGAAGGCGCGCGCAAGAGGCGCTTCGACGCGAGCTGGGCCCTGCGTGTG	960
QY	1132	GACCCACGCGCCCTGTGCGCTTCATCGCGCGCTTCGAGGACGAGAGGTTGGAC	1191
Db	961	GACCCACGCGCCCTGTGCGCTTCATCGCGCGCTTCGAGGACGAGAGGTTGGAC	1020
QY	1192	ATCATCTGSCGCCCTTGCCAAAGATCTTGGCCACCCCAAGGTGCAGATCGGCATCCTG	1251
Db	1021	ATCATCTGSCGCCCTTGCCAAAGATCTTGGCCACCCCAAGGTGCAGATCGGCATCCTG	1080
QY	1252	GGTACCGGCAAGCGCGCTACGAGAGCTGTTGAACGCCATCGGCACCAAGTACAGGGC	1311
Db	1081	GGTACCGGCAAGCGCGCTACGAGAGCTGTTGAACGCCATCGGCACCAAGTACAGGGC	1140
QY	1312	CGCGCCAAGGGCGTGTCAAGTTCTCGGCGCCCTTGGCGCACATGTTCAAGCGCGCGC	1371
Db	1141	CGCGCCAAGGGCGTGTCAAGTTCTCGGCGCCCTTGGCGCACATGTTCAAGCGCGCGC	1200
QY	1372	GACTTCATGCTGTGGCTTCGAGCGCTTCGAGCGCTTCGAGCTGCACACCGTCAAGGAG	1431
Db	1201	GACTTCATGCTGTGGCTTCGAGCGCTTCGAGCGCTTCGAGCTGCACACCGTCAAGGAG	1320
QY	1432	CACTACGGTACCGTGGCGGTGTAGCTCCACGGCGGCTTCGAGCGCTTCGAGCTGCACACCGTCAAGGAG	1491
Db	1261	CACTACGGTACCGTGGCGGTGTAGCTCCACGGCGGCTTCGAGCGCTTCGAGCTGCACACCGTCAAGGAG	1320
QY	1492	GGCGTACCGGCTTCACATPGGCGGCCCTGAACCCCGACAAGCTGCACGAGCTGACGCC	1551

Db	1321	GGCGTACCGGCTTCACATGGGGCCCTGAACCCCGACAAGCTGGACGAGGTGAGCC	1380
Qy	1552	GAGCCCTGGCCCGCACCGTGGCCGCTGCCAGGAGGTGTTCGGGCGCGCTACCCC	1611
Db	1381	GAGCCCTGGCCCGCACCGTGGCCGCTGCCAGGAGGTGTTCGGGCGCGCTACCCC	1440
Qy	1612	GAGATGTTGGCCAACTGCATCAGCCAGGACCTGTCTGTGCCAAGCCCGCCAGAGTGG	1671
Db	1441	GAGATGTTGGCCAACTGCATCAGCCAGGACCTGTCTGTGCCAAGCCCGCCAGAGTGG	1500
Qy	1672	GAGGGCCTGCTGGAGGAGTGGTGTACGGCAAGGGCGCGCTGGCCACCGCCAAAGAGGAG	1731
Db	1501	GAGGGCCTGCTGGAGGAGTGGTGTACGGCAAGGGCGCGTGGCCACCGCCAGAGGAG	1560
Qy	1732	GAGATCAAGTGGCCGTGTGCCGAGAGATCCCC	1764
Db	1561	GAGATCAAGTGGCCGTGTGCCGAGAGATCCCC	1593

```

RESULT 5
US-09-980-771A-6
; Sequence 6, Application US/09980771A
; GENERAL INFORMATION:
; APPLICANT: CNRS
; TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF 1
; TITLE OF INVENTION: METHOD OF OBTAINING THEM AND THEIR USES
; FILE REFERENCE: WO/99 AB CNR AMYL
; CURRENT APPLICATION NUMBER: US/09/980,771A
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: FR 99/06494
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: fragment of the complete cDNA coding for the GBSSI of Chi
; NAME/KEY: CDS
; LOCATION: (1)..(1314)
; OTHER INFORMATION:
; US-09-980-771A-6

```

Query Match	61.9%;	Score 1314;	DB 37;	Length 1314;
Best Local Similarity	100.0%;	Pred. No. 1.9e-177;		
Matches 1314;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

QY	172	GC	GCTGGACATCGT	GATGGTTC	GCTGAGTGC	CGCCCTTG	TCCAAAGAGCGGCGCC	CTG	231
Db	1	GC	GCTGGACATCGT	GATGGTTC	GCTGAGTGC	CGCCCTTG	TCCAAAGAGCGGCGCC	CTG	60
QY	232	GG	CGATGTGACT	GGTGGCCCTG	CCCTATT	TGAGTGGT	CAAGCGGCGCACCGGCTCAT	GACC	291
Db	61	GG	CGATGTGACT	GGTGGCCCTG	CCCTATT	TGAGTGGT	CAAGCGGCGCACCGGCTCAT	GACC	120
QY	292	ATT	GCCCTCGCT	TACGACCA	GATGAC	GCCTGGG	CACCTCGGTGCTGTGGACATC	351	
Db	121	ATT	GCCCTCGCT	TACGACCA	GATGAC	GCCTGGG	CACCTCGGTGCTGTGGACATC	180	
QY	352	AT	GGGCGAAGT	CGCTACTT	CCACTCC	ATCAAGAAGG	CGCTGCACCGCTGTGGATT	411	
Db	181	AT	GGGCGAAGT	CGCTACTT	CCACTCC	ATCAAGAAGG	CGCTGCACCGCTGTGGATT	240	
QY	412	GAC	CACCCCTG	GTTCCTG	CGCAAGG	TCTGGG	CGAACACCGGCTCCAAAGCTGTACG	CCCC	471
Db	241	GAC	CACCCCTG	GTTCCTG	CGCAAGG	TCTGGG	CGAACACCGGCTCCAAAGCTGTACG	CCCC	300
QY	472	CG	CTCGGCGCT	GACTAC	CTGGAC	AACAC	AAGCGTTTCGCCCTGTTC	GCAAGGCGCT	531
Db	301	CG	CTCGGCGCT	GACTAC	CTGGAC	AACAC	AAGCGTTTCGCCCTGTTC	GCAAGGCGCT	360

QY	343	GTGGCATACTGG-----GCGAGAAGTCCGCTACTTCACCTCCATCFAGAAG	390
Db	520	TCCGAGATCAAGGTCTGTGACAAAGTACGAGAGGGTGAAGTACTTCACCTGCTACAGCGC	579
QY	391	GGCGTCACCGGTGTGGATTGACACACCCCTGGTTCCTGSCCAAGTCTTGGGCGAAGCC	450
Db	580	GGGGTGAACCGGTGTCTGACACACCCGTGCTCTTGGAAGAGTCTCGGGGCAAGACC	639
QY	451	GGCTCCAAAGTGTACCGCCCCGCTCCGGCGCTGACTACTGTGACAAACACCAAGCGCTTC	510
Db	640	AAGSAGAAGATCTACGGGCCCGAGCGCGCAGGACTACGAGGACAAACACGACAGCGCTTC	699
QY	511	GCCTGTGTGCAAGCGCGCTATTGAGGTGTCGCCGGTGTGCCCTTTCG-----	559
Db	700	AGCCTTCTCTGCCAGCGCGCTGGAAGTCCGAGGATCCTGAACCTGCACATAACCC	759
QY	560	-----GCCCGCGGAGGACTCGCTTCGTGSCCAACGACTGGCACTCCGCG	606
Db	760	TACTTTTCTGGGCCCTACGGGGAGGAGTGGTGTTCGTGTCGANTGACTGGCACAGGGC	819
QY	607	CTGGTGCCTCTCTGTGAAGGACGAGTACACGCCCAAGSGCCAGTTCACCAAGGCCAAG	666
Db	820	CTTCTGCCCTGTACTCTCAAGAGCAACTACCAAGTCCAAATGGCATCTACAGGGCCGCAAG	879
QY	667	TCGGTGTCTGCTATCCACAACATCGCTTCCAGGGCGCATGTGGGAGAGGCTTTCAG	726
Db	880	GTGGCATCTGCAATCCACAACATCTGTAACAGGGCGCTTCTCTTCGACGACTTCGCG	939
QY	727	GACACGAAGCTGCCAGCGCGCTTTGACAAGCTGGCTTCTCGGACGCTATGCCAAG	786
Db	940	CAGCTCAACTCGCCGACAGGTTCAAGTGTCTTCGACTTCAATCGACGGCTACGACAAG	999
QY	787	GTTTACACTGAGGCCACCCCATGAGGAGGACGAGAAGCCCGCTGTACGGGAAAGACC	846
Db	1000	-----CCGGTGGAGGGCGC-----	1014
QY	847	TACAAGAGATCAACTGGCTGAAGGTGGCATATTTCGCCGCCACAAAGCTGTGACTGTG	906
Db	1015	-----AAGATCTGATTTGAAGCGCGGATCCTGCAAGCCGACAAAGTGTCTGACGGTG	1068
QY	907	TCGCCCAACTACGCGACCGAGATCGCTGCCGATGTCGCGCGGTGTGGAGCTGGACACC	966
Db	1069	AGCCCTTACTACGGGAGGASCTCATCTCTGGGGAAGCCAGGGCTGCGAGCTCGACAAC	1128
QY	967	GTCACTCGCCCAAGGGCATTTGAGGGCAITTTGAACGGCATTGACATTTGAGGAGTGGAA	1026
Db	1129	ATCATGCGCTCACTGGGATCACCGGCATCGTCAACGGCATGATGTTAGCGAGTGGAC	1188
QY	1027	CCCAAGACCAAGATTCTCTGTCTGCCCTTACACCAAGACAGGCTCTACGCCGGGAAG	1086
Db	1189	CCCAACCAAGACAAGTTCTTCGCGCTCAACTAGACATCACCCCGCTTTGAGGGGAAG	1248
QY	1087	GCGCCGCCAAGGAGGCCCTGCAGGCGAGCTGGCGCTGCTCTGGACCCACCGCCGCC	1146
Db	1249	CGGCTGAACGAAGAGGCGCTGCAGGCGAGGTGGGGCTGCCGTGGACCGGAAGTGGCC	1308
QY	1147	CTGTTCGCTTCATCGCCCGCTCGGAGSAGCAAGGGTGTGGACATCATCTCTGGCGCC	1206
Db	1309	CTGTGGCGGTTCATCGCAGCTGGAGGACGAGAAGGCCCGGACGTGATGATCGCGCC	1368
QY	1207	CTGCCCAAGATCTCTGGCCACCCCAAGGTGCAGATGCCATCTCTGGTACCGCAAGGCC	1266
Db	1369	ATCCCGGAGATCTCTGAAGGAGGAGGACGTCCAGATCTTCTCTGGCACCGGGAAG	1428
QY	1267	GCCTACGAGAGCTGGTGAACGCCATCGGCACCAAGTACAAGGGCCGCGCAAGGGCGTG	1326
Db	1429	AAGTTCGAGCGGCTTACTCAAGAGCATTTGAGGAAATTCCTCGAGCAAGGTGAGGGCGTG	1488
QY	1327	GTCAAGTTCTCGGCGCCCTTGGCGCATGCTCAACCGCGCGCGCACTTTCATCTGGTG	1386
Db	1489	GTCAAGTTCAACGCGCGCTGGCTTCACCAAGATATGGCGCGCGGACGTCTCGCGCTC	1548
QY	1387	CCCTCGGCTTCGAGCCCTCGGGCTGATCCAGCTGCACGCCATGCACTACGCTACCGG	1446

```

1  RESULT 8
2  US-07-771-351-3
3  ; Sequence 3, Application US/07771351
4  ; GENERAL INFORMATION:
5  ; APPLICANT: ILLIIS, MARCELLA
6  ; APPLICANT: CHEN, CHE-HONG
7  ; APPLICANT: GROSS, PAUL R.
8  ; APPLICANT: FANG, LAWRENCE Y.
9  ; TITLE OF INVENTION: GENETIC SEQUENCES FOR ALTERING
10 ; TITLE OF INVENTION: GRANULE-BOUND STARCH SYNTHASE IN WHEAT
11 ; NUMBER OF SEQUENCES: 8
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
14 ; STREET: 5 PALO ALTO SQUARE, SUITE 400
15 ; CITY: PALO ALTO
16 ; STATE: CALIFORNIA
17 ; COUNTRY: USA
18 ; ZIP: 94306
19 ; COMPUTER READABLE FORM:
20 ; MEDIUM TYPE: Floppy disk
21 ; COMPUTER: IBM PC compatible
22 ; OPERATING SYSTEM: PC-DOS/MS-DOS
23 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
24 ; CURRENT APPLICATION NUMBER: US/07/771,351
25 ; FILING DATE: 19911004
26 ; CLASSIFICATION: 435
27 ; ATTORNEY/AGENT INFORMATION:
28 ; NAME: NEELEY, RICHARD L.
29 ; REGISTRATION NUMBER: 30092
30 ; REFERENCE/DOCKET NUMBER: LFCP-015/0005
31 ; TELECOMMUNICATION INFORMATION:
32 ; TELEPHONE: 415-494-7622
33 ; TELEFAX: 415-857-0663
34 ; INFORMATION FOR SEQ ID NO: 3:
35 ; SEQUENCE CHARACTERISTICS:
36 ; LENGTH: 2279 base pairs
37 ; TYPE: NUCLEIC ACID
38 ; STRANDEDNESS: single
39 ; TOPOLOGY: linear
40 ; MOLECULE TYPE: cDNA
41 ; ORIGINAL SOURCE:
42 ; STRAIN: WHEAT VAR. PROBAND755
43 ; IMMEDIATE SOURCE:
44 ; CLONE: pSOG188

```


US-07-771-351-3

Query Match	24.7%	Score 523.8	DB 3	Length 2279	
Best Local Similarity	61.2%	Pred. No. 3e-65			
Matches	986	Conservative 0	Mismatches 532	Indels 93	Gaps 5
QY	163	GCACCTGCGCGTGACATCGTATGGTTGCTGCTGAGGTGCGCCCTTGTGTCAAAGACG	222		
Db	391	GGCAGCGCGCATGAACCTCGTGTTCGCGCCGAGATGGCGCCTTGAGCAAGACT	450		
QY	223	GGCGGCTGGCGGTGACTGTGTGGCTGCCTATTGAGCTGGTCAAAGCGCGCACCGC	282		
Db	451	GGCGGCTGGCGACGTCTCGGGGGCTCCCGCCGCAATGGCGCCACGGTACCGG	510		
QY	283	GTCAATGACCAATGCCCTCGTAGCAGCAGTACGCTGACGCTGCGGACACCTCGGTGTGTC	342		
Db	511	GTCAATGCTATCTCCCGCGCTACGACCACTACAAGCAGCGCTGGGACACCAAGCGTATC	570		
QY	343	GTGACATCATGG-----GCGAGAAGTCCGCTACTTCCACTCCATCAAGAAG	390		
Db	571	TCCGAGATCAAGTCTGTTGACAGTACGAGAGGTGAGGTACTTCCACTGCTACAGCGC	630		
QY	391	GGCGTGACCGGTGTGGATTGACCAACCCCTGGTCTTGCCCAAGTCTTGGGGCAAGACC	450		
Db	631	GGGGTGACCGCGTGTGTCGACCAACCCGTGCTTCTTGAGAAGTCTCGGGCAAGACC	690		
QY	451	GGCTCCAAGTGTACGGCCCGCTCGCGCGCTGACTACCTGGACCAACCAAGCGTTC	510		
Db	691	AAGGAGAAGATCTATGGACCCGACGCGGCAACGACATACGAGGACAACAGCAGCGTTC	750		
QY	511	GCCTGTCTGCAAGGCGGTATTGAGGCTGCCGCGTGTGTCGCCCTTCGCGCC-----	563		
Db	751	AGCCTTCTCTCGAGGCGAGCAGTGTGAGTGGCCAGGATCTCGACCTCAACAACAACCA	810		
QY	564	-----CGCGGAGGACTGGTCTTCGTGGCCACGACTGGCACTCCGCC	606		
Db	811	CACCTTCTGGACCTTACGGGGAAGAGCTGGTGTGTGTGCAACGACTGGCAGCAGGGC	870		
QY	607	CTGGTGGCCGTCTGCTGAAGGACGAGTACACGCCCAAGGCGCAAGTTCACCAAGGCCAAG	666		
Db	871	CTTCTGGCCTGCTACCTCAAGAGCAACTACCAAGTCCAATGSCATCTATAGGACGCCAAG	930		
QY	667	TCGGTGTGCTATCCACAACATCGCCTTCAGGGCGCGATGTGGGAGGAGCTTTCAG	726		
Db	931	GTGGCATCTGTCATCCACAACATCTGTCACGAGGCGGCTTCTCCTTCGACGACTTCGCG	990		
QY	727	GACACAAGTGCGCCACGCGCGCTTTGACAAGCTGGCCTCTCGGACGCGTATGCCAAG	786		
Db	991	CAGCTCAACCTGCCGACAGGTTCAAGTCGTCTCTCGACTTCATCGACGCTACGACAAG	1050		
QY	787	GTTTACACTGAGCCACCCCATGGAGGAGACAGAGCCCGCGCTGACGGGAAGACC	846		
Db	1051	-----CCGGTGGAGGCGCG-----	1065		
QY	847	TACAAGAAGATCACTGGCTGAAGGTGGCATTATCGCGCGACAGCTGGTGACTGTG	906		
Db	1066	-----AAGATCACTGGATGAAGCGCGGATCTTCGAGGCCACAGGTGCTGACTGTG	1119		
QY	907	TCGCCCACTACCGCACCGAGATTCGTGTCGATGCCGCGCGGTGTGGAGCTGGACAC	966		
Db	1120	AGCCCCCTACTATCTGAGGAGCTAACTCTGCGGAAGCCAGGGGCTCGAGCTCGACAAC	1179		
QY	967	GTATCCGCGCCCAAGGGCATTTAGGGCATTTGTAACGGCATGACATTTGAGGAGTGAAC	1026		
Db	1180	ATCATCGGCTCACTGGGATCACCGGCATGGTCAACGGCATGACGCTCAGCGAGTGGGAC	1239		
QY	1027	CCCAAGACGCAAGTTCCTGCTCGCGCCCTACGACCAAGAACCGCTCTACGCGGCAAG	1086		
Db	1240	CCCATCAAGSACAAGTTCTCACCGTCAACTAGACGCTCACACCGCGTTGAGGGGAAG	1299		
QY	1087	GCGCGCCCAAGAGGCGCTGACGCGGAGCTGGGCGCTGCCTGTGGACCCCAACCGCCCC	1146		
Db	1300	GGCGCTGAACAGAGCGCTGACGGCGAGTGGGCTGCGGCTGACCGCAAGGTGCCC	1359		

Qy	1147	CTGTTCCGCCTTCATCGCCGCCCTGGAGGACGAAGAAGGTGTGGACATCATCTCTGGCCGC	1206
Db	1360	CTGTGGCGTTCATCGCAGGCTGGAGGAGCAAGAAGGCCOCGACGTGATGATCGCCGCC	1419
Qy	1207	CTGCCAAGATCTTGCCACCCTCCCAAGTGCATCGCCATCTCTGGTACCGCAAGGCC	1266
Db	1420	ATCCCGAGATCTGAAGGAGGAGCGTCCAGATGTTCTCTGGCCACCGGAAAGAAG	1479
Qy	1267	GCCITACGAAGAAGTGTGAACGCGCATCTGGCAACAAGTACAAGGCCGCCCAAGGCGGTG	1326
Db	1480	AAGTTTGAGCGGCTGCTCAGACGCTGGAGGAGAAGTTCGCCACCAAGTGAAGGCCGTG	1539
Qy	1327	GTCAGTTCTGGCGCCCTCGGCACATGCTCACCGCGCGGCCACTTCATGCTGGTG	1386
Db	1540	GTCAGTTCAAACGCGCTGCTCACCAGATGATGCCGGCGCACGTGCTGGCGGTG	1599
Qy	1387	CCCTCGGCTTCAGCCCTCGCGGCTGATCCAGCTGCAGCCCATGCACCTACGGTACCGTG	1446
Db	1600	ACACGCGCTTCAGGCGCTCGGCGCTCATCCAGTCCAGGGAATGCGTACGAAACGCCG	1659
Qy	1447	CCGTGGTAGCCTCCAACCGCGGCTTGGTCGACACCGTCAAGAGGGCGCTCACCGGCTTC	1506
Db	1660	TGCGCTTCGCGTCGACAGCGGCTCTGCGACACTATCTGGAAGCAAGACCGGTTCT	1719
Qy	1507	CACATGGCGCCCTGA-----ACCCGACACAGCTGGAGAGGCTGACGCCGACGCCGTG	1560
Db	1720	CACATGGCGGCTTCAGCGTTGACTCAACGTTGTTGGAGCCGCCGACGTGAAGAAGGTG	1779
Qy	1561	GCCGCCACGCTGGCGTGCACGAGGTGTGTGCGGGCGCGCTACCCGATGGTG	1620
Db	1780	GTACACACCTGAAGCGTGCCGTCAAGTCTGTCGGCACCGCGCGGTACCATGAGTGGTTC	1839
Qy	1621	GCCAACTGCATCAGCCAGGACCTGTCTCTGGTCTCAAGCCCGCCGACGAAGTGGGAGGCGCTG	1680
Db	1840	AAGAACTGCATGATACAGGATCTCTCTGGAGGGGCTTSCCAAGNACTGGGAGGACGTG	1899
Qy	1681	CTGAGGAGGTGTGTGTCAGGCAAGGCGCGGTGGCCACCGCCCAAGAAGAG	1731
Db	1900	CTTCTGGAACTGGGGTGGAGGGAGCGAGCAGGCGATCTGTGGCGGAGAG	1950

```

RESULT 9
US-09-731-166-3
; Sequence 3, Application US/09731166
; GENERAL INFORMATION:
; APPLICANT: Sewall, Vincent J. H.
; APPLICANT: Singletary, George W.
; TITLE OF INVENTION: Production Of Modified Polysaccharides
; FILE REFERENCE: 35718/206348
; CURRENT APPLICATION NUMBER: US/09/731,166
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/169,993
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: UDP-glucose starch glycosyl transferase
; OTHER INFORMATION: Accession No. 22509
; NAME/KEY: CDS
; LOCATION: (1)...(1818)
US-09-731-166-3

```

Query Match 22.5%; Score 478; DB 29; Length 1818;
Best Local Similarity 59.8%; Pred. No. 1e-58;
Matches 986; Conservative 0; Mismatches 565; Indels 99


```

: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FASTSEQ for Windows 2.0
: CURRENT APPLICATION DATA:
: FILING DATE: September 2, 1994
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA: None
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 208/153
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2267 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-300-726A-1

```

Query Match 22.58; Score 478; DB 7; Length 2267;

Best Local Similarity 59.88; Pred. No. 9.7e-59;

Matches 986; Conservative 0; Mismatches 565; Indels 99; Gaps 7;

```

QY 161 GTGCGGCTTCGGCGTGGACATGATGTTGCTGTGAGTGGCGCCCTTGTPCCAAAGA 220
DB 403 GCGCCAGCGCGGATGAAGCTGTCTGTCGGCGCGGAGATGGCCCGTGGAGCAAGA 462
QY 221 CGGGCGGCGTGGCGATGTGACTGGTGGCGCTGCTATTGAGCTGGTCAAGCGCGCCACC 280
DB 463 CGGGCGGCGTGGCGAGCTCTCGGCGGCGCTGCGCGGCGCATGCGCGCGAATGGGCACC 522
QY 281 GGTCTAGACCATGCCCCCTGCTAGCAGCAGTACGCTGAGCGCTGGGACACCTCGGTGG 340
DB 523 GTGTCTATGTTGCTCTCTCCCGGCTTACGACCAAGTACAGGAGCGCTGGGACACCGCTG 582
QY 341 TCGTGGACAT-----CATGGCGGAGAGGTCCGCTACTTCCACTCCATCAAGA 388
DB 583 TGTCGAGATCAAGATGGGAGACAGGTACGAGCGGTCAAGTCTTCCACTGCTACAGC 642
QY 389 AGGGGTGACCGCGTGTGATTTGACCAACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGA 448
DB 643 GGGAGTGGACCGGCTTCTGTTGACCAACCCACTGTTCTCTGGAGAGGTTTGGGAAAGA 702
QY 449 CCGGCTCCAAAGCTGTAGCGGCCCGGCTCGGCGCTGACTACCTGGACCAACCAAGCGCT 508
DB 703 CCGAGGAGAGATCTAGGGCGCTGACCTGGAGCGGACTACAGGGAACCAAGCTGGGT 762
QY 509 TCGGCTGTGTGCAAGCGCGCTATTTGAGGTGCGCGGCTGCTCCCTTCGCG----- 561
DB 763 TCAGCCTGCTATGCCAGGACGACTTCAAGCTCAAGGATCCTGAGCCTCAACAACAACC 822
QY 562 -----CCCGGCGAGGACTCGGTCTTCTGGTGGCCAAAGTCTGACATCGG 604
DB 823 CATACTTCTCCGACCATACAGGGGAGGACGCTGTTCTGCTTGTGCAACGACTGGCACACC 882
QY 605 CCCTGGTGGCGCTCTGCTGAGGACGAGTACCAAGCCCAAGGGCCGCTTACCAAGGCCA 664
DB 883 GCGCTCTCTGCTACTTCAAGAGCAACTACCAAGTCCCAAGGCGATCTACAGGAGCAAA 942
QY 665 AGTCGGTGTGGCTGTATCCAAACATCGCTTCCAGGGCGCGCATGTGGGAGGAGGCTTCA 724

```

```

DB 943 AGACCGGCTTTCGATCCACAACAATCTCTTACAGGGCGCGGTTTCGCCCTTCTCCACATACC 1002
QY 725 AGGACACGAAGTGCCTCCCGAGCGCCCTTTTGACAAAGCTTGGCCCTTCTCGAGACGGCTATGCCA 784
DB 1003 CGGAGCTGAACCTCCCGAGAGATTCAAGTCGTCCTTCGATTTCATCGACGGCTACGAGA 1062
QY 785 AGGTTTACATGAGGCCACCCCATGGAGGAGGACGAGAGCCGCCCTGACGGGAAAGA 844
DB 1063 AG-----CCCGTGAAG----- 1074
QY 845 CCTACAAAGATCAACTGCTGAAGGTGCTGATTCGCGCGCGCAAGCTGGTGAAGT 904
DB 1075 ---GCCGGAAGATCAACTGGATGAAGCGCGGATCTCTCGAGGCGGACAGGGTCTCTACCG 1131
QY 905 TGTGCGCCCACTAGCGGACGAGATCGTCCGATGCGCGGCGGTGTGAGCTGGACA 964
DB 1132 TCAGCCCTTACTAGCGCGGAGGAGTCTCTCCGCGCATCGCAGGGGTGGAGCTGCACA 1191
QY 965 CCGTCATCCGGGCAAGGGCATTTGAGGCGATTGTGAGGGCATGAGCATTTGAGAGTGA 1024
DB 1192 ACATCATCGGCTTCCCGGATCACCGGCTCGTCAACGGCATGCGATCAGCGTACGAGTGG 1251
QY 1025 ACCCAAGACCGACAAGTTCCTGCTGCGCCCTAGCAGCAGACAGAGCTCTACCGCGCA 1084
DB 1252 ACCCAGGAGGACAGTACATCGCGTGAAGTACGAGTGTGACGCGCTGGAGGCCA 1311
QY 1085 AGCGCGCGCAAGGAGGCGCTGCGAGCGGCTGCGCTGCTGAGACCCACCGCC 1144
DB 1312 AGCGCTGAACAAGGAGGCGCTGCGAGCGGAGTGGCTCCCGTGCAGCGGAATCC 1371
QY 1145 CCCTGTCGCTTCATCGCGCGCTGGAGGAGAGAGGTTGTGAGCATCATCTCTGGCG 1204
DB 1372 CGCTGGTGGGCTTTCATCGGAGGCTGGAAGAGAGAGGACCGGCTCATTTGCGCGCG 1431
QY 1205 CCCTGCCCAAGATCTGGCCACCCCAAG---GTGCAGATCGCCATCTGGTACCAGCA 1261
DB 1432 CCATCCGCGACTCATGGAGATGTTGGAGGACGTGACAGTCTGTTCTGGGCGACGGCA 1491
QY 1262 AGCGCGCTACGAGAAGCTGGTGAACGCCATCGGCACCAAGTACAAAGGCGCGCCCAAG 1321
DB 1492 AGAAGAGTTCGAGCGCATGCTCATGAGCGCGGAGGAGAGTTCAGGCAAGGTGGCG 1551
QY 1322 GCTGTGCTCAAGTTCGCGCGCTTGGCGCATGCTCACCGCGCGCGCCGCTCATGCG 1381
DB 1552 CCGTGGTCAAGTTCACCGCGCTGCGCGCACCATCATGCGCGCGCGCGCGCTGCTG 1611
QY 1382 TGTGCGCTTCGCGCTTCGAGCGCTGCGCGCTGATCCAGCTGACGCGCATGCTACGCTA 1441
DB 1612 CCGTCACGCGCTTCGAGCGCTTCGCGCTTCATCCAGCTGCGAGGGATGCGATACGAA 1671
QY 1442 CCGTGGCGCTGGTACGCTTCCACCGCGCGCTGGTTCGACACCGTCAAGGAGGCGCTACCG 1501
DB 1672 CGCCTGCGCTGCGCTTCCACCGGTGACTGCTGACACCATCATGAGGCAAGACG 1731
QY 1502 GCTTCCACATGGCGCGCTGA-----ACCCCGACAAGCTGGAGGCTCACCGCGAG 1555
DB 1732 GGTTCACATGGCGCGCTTCAGCGTCAAGTCTGAGGCGCGCGCGCGCTCAAGA 1791
QY 1556 CCGTGGCGCGCACCGTGGCGCTGCGAGCGGCTTTCGCGGGCGCGCTTACCGCGAGA 1615
DB 1792 AGGTGGCACACCTTGCAGCGCGCATCAAGTGGTGGCGACCGCGCGCTTACGAGAGA 1851
QY 1616 TGGTGGCAACTGATCAGCCAGGACCTGCTGCTGCTCCAGCCCGCGCGCAAGTGGGAGG 1675
DB 1852 TGGTGAAGAACTGATCCAGGATCTCTCTGGAAGGGCGCTTCCAAAGAACTGGGAGA 1911
QY 1676 GCCTGCTGGAGGAGTGTGTACGGCAAGGCGCGCTGGCCACCGCGCAAGAGAGGAGA 1735
DB 1912 ACCTGCTGCTACGCTCGGGTCCGCG---GGCGCGAGCCAGGGGTCTGAAGGCGAGAGA 1968
QY 1736 TCAAGGTGCCGTTGCCGAGAAAGATCCCGC 1765
DB 1969 TCGCGCGCTCGCCCAAGGAGAACTGGCGC 1998

```


1384 GTGCGCTCGCTTACAGCCCTGGCGCTGATGATCCAGCTGCACGCCATGACACATACGATACC 1443
1426 GTTACACAGCGCTTGCAGCCCTGGCGCTTATCCAGCTGCAGGGGATGCGATACGGAACG 1485
1444 GTGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1503
1486 CCTGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1545
1504 TTCCACATGGCGCTTGAACCCG-----ACAAGCTGGAGGAGGCTGACCGCGACGCC 1557
1546 TTCCACATGGCGCTTGAACCCG-----ACAAGCTGGAGGAGGCTGACCGCGACGCC 1605
1558 CTGCGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1617
1606 GTGCGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1665
1618 GTGCGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1677
1666 GTGCGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1725
1678 GTGCGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1737
1726 GTGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1782
1738 AGGTGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1765
1783 GCGCGCTGATGATCCACCGCGCTGATGATCCAGCTGCAGGGGATGCGATACGGAACG 1810

RESULT 14

US-09-654-617-385561
; Sequence 385561, Application US/0954617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 385561
; LENGTH: 2566
; TYPE: DNA
; ORGANISM: Oryza sativa
; OTHER INFORMATION: unsure at all n locations
US-09-654-617-385561

Query Match 21.1%; Score 449.2; DB 25; Length 2566;
Best Local Similarity 58.6%; Pred. No. 1.2e-54;
Matches 933; Conservative 0; Mismatches 563; Indels 96; Gaps 5;
171 CGCGTGGACATCGTATGATGCTGCTGAGTGCAGCTGCGCCCTTGTGTCACAGCGCGCGCT 230
695 CGCGATGAACGCTGCTGCTGCGCGCGAGATGCGCCCTTGAAGCAAGACCGCGCGCT 754
231 GCGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290
755 CGGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
291 CATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350
815 GATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874
351 CATGG-----GCGAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 398
875 CAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934
399 CGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
935 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994
459 GCTGTACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 518

995 GATCTAGGACCTGACACATGAGCTGATGATTAACAAGAACACAGATGCTTACGCTTCT 1054
519 CTCAAGAGCGCTTATGAGGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
1055 TTGCGAGGACGACTGAGGCTCTAGGATCTTAACCTCAACAACCAACCAATACATTCAA 1114
565 -----GGCGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
1115 AGGAACCTTATGCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1174
615 CGTCTCTGCTGAAGGACGAGTACAGCCCAAGGCGCAGTTTCAACAAGGCGCAGTTCGCTGCT 674
1175 GAGCTACCTGAAGAACAACCTACAGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
675 GGTATATCCAAACATCGCTTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
1235 CTGCTATCCAAACATCTCTTACAGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1294
735 GCTGCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794
1295 CTTCTCGAGAGGTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
795 TGAGGCGCACCCCATGAGGAGGACGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
1344 -----ACCGCGTGGAGGCGAGAA 1363
855 GATCAACTGCTGAAGGCTGGCATTTATCGCGCGCGACAAAGCTGCTGCTGCTGCTGCTGCT 914
1364 GATCAACTGCTGAAGGCTGGCATTTATCGCGCGCGACAAAGCTGCTGCTGCTGCTGCTGCT 1423
915 CTACGCGACGAGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974
1424 CTACGCGAGGAGCTCATCTCCGCGCATCGCCAGGAGTGGAGCTGCTGCTGCTGCTGCTGCT 1483
975 CGCAGAGGCTTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
1484 GTCACCGGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543
1035 CGACAAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094
1544 GGACAAAGTATCATCCGCGCAAGTACGAGCAACACCGCAATCGAGCGAGCGCTGCTGCT 1603
1095 CAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1154
1604 CAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1663
1155 CTTATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1214
1664 GTTCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723
1215 GATCTGCGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274
1724 GCTCATG---CAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1780
1275 GAAGCTGCTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1334
1781 GAAGCTGCTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840
1335 CTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1394
1841 CRACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1900
1395 CTTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454
1901 CTTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1960
1455 AGCTTCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1514
1961 CGCGTCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2020
1515 CGCGTCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1568
2021 CGGTCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2080

Db 2201 CCTGGGCGCTGCGCGCGCAGCGCGCGGGGATCG 2232
||| || | ||| || | ||
||| || | ||| || | ||

Search completed: June 7, 2003, 01:55:08
Job time : 3358.08 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:50:01 ; Search time 982.928 Seconds

(without alignments)
13677.072 Million cell updates/sec

Title: US-09-980-771A-2

Perfect score: 2124

Sequence: 1 atggctgtgctctaccag.....ccgcaccgcgaagtcgcgc 2124

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7125412 seqs, 3164683403 residues

Total number of hits satisfying chosen parameters: 14250824

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2.6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2.6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2.6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2.6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2.6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2.6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 7: /cgn2.6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 8: /cgn2.6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 9: /cgn2.6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 10: /cgn2.6/ptodata/2/pna/US60_NEW_COMB.seq.*
- 11: /cgn2.6/ptodata/2/pna/US60_NEW_COMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478	22.5	1818	1	PCT-US02-27129-7
2	478	22.5	1818	9	US-10-228-063-7
3	478	22.5	2263	1	PCT-US02-33122-1
4	478	22.5	2263	9	US-10-272-291-1
5	476.4	22.4	2263	1	PCT-US02-33122-2
6	476.4	22.4	2263	9	US-10-272-291-2
7	472.8	22.3	1915	9	US-10-336-753-46
8	331	15.6	2223	1	PCT-US02-27129-9
9	331	15.6	2223	9	US-10-228-063-9
10	305.2	14.4	2098	8	US-10-425-114-4084
11	300.6	14.2	1416	8	US-10-425-114-19555
12	298	14.0	1406	8	US-10-425-114-17713
13	298	14.0	1409	8	US-10-425-114-16641
14	277.8	13.1	3290	8	US-10-424-599-64934
15	219.2	10.3	2423	9	US-10-336-753-48
16	219.2	10.3	2480	9	US-10-109-048-1144
17	213.2	10.0	1502	8	US-10-425-114-31744
18	213.2	10.0	1828	8	US-10-425-114-1052
19	207	9.7	2010	9	US-10-336-753-50
20	207	9.7	2865	9	US-10-109-048-1143
21	206.8	9.7	1888	8	US-10-425-114-34283
22	204	9.6	1286	8	US-10-425-114-8641

23	202.8	9.5	1428	8	US-10-369-493-44118	Sequence 44118, A
24	199.8	9.4	1623	8	US-10-425-114-2836	Sequence 2836, Ap
25	198.4	9.3	618	7	US-09-837-604A-73243	Sequence 73243, A
26	196.4	9.2	1377	8	US-10-369-493-42867	Sequence 42867, A
27	188.2	8.9	992	8	US-10-425-114-15417	Sequence 15417, A
28	188.2	8.9	1386	8	US-10-369-493-39259	Sequence 39259, A
29	171.2	8.1	1386	8	US-10-369-493-39632	Sequence 39632, A
30	171.2	8.1	1386	8	US-10-369-493-40000	Sequence 40000, A
31	169.4	8.0	536	7	US-09-837-604A-75067	Sequence 75067, A
32	154.8	7.3	1383	8	US-10-369-493-44603	Sequence 44603, A
33	154.4	7.3	1417	8	US-10-424-599-56423	Sequence 56423, A
34	152.6	7.2	588	5	US-09-531-113-10143	Sequence 10143, A
35	152.6	7.2	588	6	US-09-531-113-10143	Sequence 10143, A
36	152.4	7.2	532	8	US-10-401-229-4112	Sequence 4112, Ap
37	149.6	7.0	1003	8	US-10-424-599-56421	Sequence 56421, A
38	144.6	6.8	1443	8	US-10-369-493-28518	Sequence 28518, A
39	144.6	6.8	1452	8	US-10-369-493-31277	Sequence 31277, A
40	143.2	6.7	1440	8	US-10-369-493-35062	Sequence 35062, A
41	143.2	6.7	1440	8	US-10-369-493-38466	Sequence 38466, A
42	143.2	6.7	1440	8	US-10-369-493-38620	Sequence 38620, A
43	143.2	6.7	1440	8	US-10-369-493-38957	Sequence 38957, A
44	140.8	6.6	1248	8	US-10-446-203-3687	Sequence 3687, Ap
45	140.8	6.6	1536	8	US-10-446-203-3862	Sequence 3862, Ap

ALIGNMENTS

RESULT 1

PCT-US02-27129-7
; Sequence 7, Application PC/TUS0227129
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: PCT/US02/27129
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
PCT-US02-27129-7

Query Match					22.58; Score 478; DB 1; Length 1818;
Best Local Similarity					59.8%; Pred. No. 1e-56;
Matches					986; Conservative 0; Mismatches 565; Indels 99; Gaps 7;
QY	161	GTGCCACTTCGGCGCTGGACATCGTGTGCTGCTGAGTGGCGCCCTTGGTCCAGA	220		
DB	218	GCGCCAGCGCGCATGAACGTCTCTCGCGCGGAGATGGCCCTGGAGCAAGA	277		
QY	221	CGGGCGGCTGGCGGATGTGACTGGTGGCTGCTATGAGTGGTCAAGCGGGCAAC	280		
DB	278	CCGGAGGCTTCGGCGAGCTCTCGGGCGCTCGCGCGGCGGCGGCGGCGGCAAC	337		
QY	281	GCCTCATGACATGCCCCCTCGCTAGCAGCAGTACGCTGAGCGCTGGGACACCTCGGTGG	340		
DB	338	GTGTCATGTCGTCTCTCCCGCTAGCAGGACACAGGAGCGCTGGGACACGAGGTGCG	397		
QY	341	TCGTGACATCA-----TGGGCGGAGAGTCCGCTACTTCCATCCATCAAGA	388		
DB	398	TGTCGAGATCAAGATGGGAGCGGTACGAGCGGTACAGGTCTTCCACTGCTACAGC	457		
QY	389	AGGGCGTGCACCGCTGTGATGACACCCCTGGTTCCTGGCCAAAGTCTTGGGGCAAGA	448		
DB	458	CGGAGTGGACCGGTGTTCGTTGACCCACCCCTGTTCTCTGGAGAGGGTTTGGGGAAGA	517		
QY	449	CCGGCTCCAAAGCTGTAGGCGCCCGCTCGGCGCTCACTACCTGGGACACCAAGCGCT	508		

Db	518	CCGAGGAGAGATCTACGGGCTGTCTCGCTGGAACGACTACAGGGACAACCACTGGGGT	577
QY	509	TCGCCCTGTCTTCAAGAGCGCTATTAGGCTCCCGCGTGTCTGCCCTTCGGC-----	561
Db	578	TCAGCCTGTCTATGCCAGGACGACTTGAAGCTCCAAGGATCCTGAGCTTCAACAACAAC	637
QY	562	-----CCGGCGGAGACTGGCTTCTCTGGGCCCAAGCACTGSCACTCCG	604
Db	638	CATACCTTCGCGAACCATACGGGAGAGACTGCTGTCTCGTCTCAACGACTGCACACCG	697
QY	605	CCCTGTGTGCCGCTCTGCTCAAGGAGAGTACAGGCCCAAGGGCCAGTTCACCAAGGCCA	664
Db	698	GCCTCTCTGCTACTCTCAAGAGCAACTACAGTCCACGCGCATCTACAGGAGCGCAA	757
QY	665	AGTCGGTGTGGCTATCCACAACATGCCCTTCAGAGGCCGCACTGTGGGAGGAGCTTTCA	724
Db	758	AGACCGCTTTCTGCATCCACAACATCTCTTACCAAGGCGGTTCCGCTTCTCGGACTAAC	817
QY	725	AGGACAGGAAGCTGCCCCAGCGCGCTTTGACAAGCTGGCTTCTCGGACGGCTATGCCA	784
Db	818	CGGAGCTGAACCTCCCGAGAGATTCAAGTCGTCTTCGATTTCATCGACGGTACGAGA	877
QY	785	AGGTTTACACTGAGGCCACCCCATGAGGAGGACGAGAAGCCCGCTGACGGGAAGA	844
Db	878	AG-----CCGCTGGAAGG-----	890
QY	845	CCTACAGAAGATCAACTGGCTGAAGGGTGGCATTTATCGCGCGGACAAGCTGTGACTG	904
Db	891	----CCGGAAGATCACTGGATGAAGCCGGGATCTCTGAGGCCGACAGGGTCTCTACCG	946
QY	905	TGTGCGCCAACTACGGAGCCGAGATCGCTGCCGATGCCCGCGGTGTGGAGCTGGACA	964
Db	947	TCAGCCCTACTACGCGAGGAGTCACTCCGCGCATCGCAGGGGCTGCGAGTCGACA	1006
QY	965	CCGTCATCCGCGCAAGGGCATTCAGGGCATTTGTACGCGCATGGACATTCAGGAGTGA	1024
Db	1007	ACATCTCGCGCTCACCGGCATCACCGCATCGTCAACGGCATGGAGCTCAGCGAGTGG	1066
QY	1025	ACCCAGAACCGCAAGTTCTCTCTGCGCCCTACGACCAGAACAGCGTCTACGCGCGCA	1084
Db	1067	ACCCAGCAGGACAAGTACATCGCGTGAATACGACGTGTCGAGGCCGTGGAGGCCA	1126
QY	1085	AGCGCGCGCCAGGAGGGCCCTCGAGCCGAGCTGGGCGTCTGTGGACCCCAACCGCC	1144
Db	1127	AGCGCTGGAACAGGAGCGCTCGAGCGGAGTTCGGGCTCCGCGGTGGACCGGAACTOC	1186
QY	1145	CCCTGTTCGCTTCACTCGCGCGCTCGAGAGCAGAAAGGTGTGGACATCATCTGSCCG	1204
Db	1187	CGTGTGGGTTTCATCGGCAGGCTGGAAGAGCAGAGGGCCCCGAGCTCATGGCGCG	1246
QY	1205	CCCTGCCCAAGATCTTGGCCACCCCAAG---GTGCAGATTCGCATCTCTGGGTACCGGCA	1261
Db	1247	CCATCCCGCAGCTCATGGAGATGGTGAGAGACGTGCAGATCGTTCTGCTGGCACAGGCA	1306
QY	1262	AGGCGCGCTACGAGAAGCTGTCAACGCCATCGSCACCAGTACAGAGGCCGCGCCCAAGG	1321
Db	1307	AGAAGAGTTTCGAGCGCATCTCATGAGCGCCGAGAGAGTTCCTCCAGGCAAGTGC	1366
QY	1322	CGTGTGTCAAGTTCTCGCGCCCTCGGCGCACATGTCTACCGCGCGCGCGCACTTCATGC	1381
Db	1367	CCGTGTCAAGTTCAAGCGGCGCTGGCGCAACCATCATGTCGCGGCGCGGAGTGTCTCG	1426
QY	1382	TGTTGCCCTCGGCTTCGAGCCCTCGGGCTGATCCAGTGCAGCGCATGCATACGGTA	1441
Db	1427	CCGTTCACACCGGCTTCGAGCCTGGGGCTCATCAGTGTGCAAGGGGATCGCATACGGAA	1486
QY	1442	CCGTGCGCGTGTAGCTCCACCGGGCGCTGTGTCACACCGTCAAGAGGGGCGTCCACCG	1501
Db	1487	CGCCCTCGGCTCGCGGTCCACCGGTGGACTGTCTGCACCATCATCAAGGCAAGACCG	1546
QY	1502	GCTTCCACATGGGCGCCCTGA-----ACCCCGACAAGCTTGGACGAGGCTGACGCGCAAG	1555
Db	1547	GGTTCCACATGGGCGGCTCAGGCTGACTGCAACGCTGTGGAGCGCGGAGCTCAAGA	1606

QY		605	CCCTGGTGC	CGCCTGCTCTGCTGAAGACGAGATACAGGCCCAAGAGGCCAGTTTCACCAAGGCCA	664
Db		698	GCCTCTCTCTCGTGCTACTCTCAAGAGAACAATACAGTCCCACGCACTCTACAGGACGCAA	757	
QY		665	AGTCGTGTC	GCGTATCCACAACATCGCTTTCAGGGCCGCATGTGGGAGGAGCGTTTCA	724
Db		758	AGACCGCTT	TTCATCATCCACAACATCTCTTACCAGGCGCGTTTCGCTTCTCTCGAATAACC	817
QY		725	AGGACAGAA	GCTGCCACAGCGCGCTTTTCACAAGCTGGCTTCTCGGACGGGTATGCCA	784
Db		818	CSGAGCTGA	ACCTCCCGGAGAGATTCAAGTCGTCTTCGATTTCTACGACGGGTACGAGA	877
QY		785	AGGTTTAC	ACTGAGGCCAACCCCATGGAGGACGAGAACGCCCGCTCACGGGAAAAGA	844
Db		878	AG-----	-----CCGCTGGAAG-----	890
QY		845	CCTACAGA	AGATCAACTGGCTGAAGGTGGCAATTATCGCCGGACACAGCTGGTGACTG	904
Db		891	-----	---CGGGAAGATCAACTGGATGAAGSCCGGATCTCGAGGCCGACAGGGTCTCACCG	946
QY		905	TGTGCGCCA	ACTACGGGACCGAGATCGCTGCCCATGCCGCGCGGTGTGGAGCTGGACA	964
Db		947	TCAGCCCC	TACTACGCCGAGGAGCTCATCTCCGGCATCGCCAGGGGCTGGAGCTCGACA	1006
QY		965	CCGTATPC	CGGCCCAAGGGCATTTAGGGCATTTGTAAACGSCATGGACATTAGAGAGTGA	1024
Db		1007	ACATATGC	CCCTTCACCGGCATCACCGGCATCGTCAACGGCATGGACGTACAGGAGTGGG	1066
QY		1025	ACCCCAAG	CCGACAAAGTTCTCTCTGTGCGCCCTACGACCAGACAGCGTCTAGCCGGCA	1084
Db		1067	ACCCACAG	AGGGACAAGTACATCCGCTGAAGTACGACGTGTCCAGCGCGCTGGAGCGCA	1126
QY		1085	AGSCCGCG	CCCAAGAGGCCCTTCAGSCCGAGCTGGCCCTGCTGTGGACCCCAACCGCC	1144
Db		1127	AGSCGTGA	AACAAGAGGCGCTCAGSCGGAGTCTGGGCTCCGCGTGGACCGGAACTCC	1186
QY		1145	CCCTGTTG	CGCTTATCGGCCCTCTGAGAGCAGAAAGGTGTGGACATCATCTGGCCG	1204
Db		1187	CGTGTGG	CGTTTATCGGCAGGTGTGAAGAGCAGAAAGGCCCCCGACGTATGGCGCGG	1246
QY		1205	CCCTGCCA	AGATCTTGGCCACCCCAAG---GTGCAGATCGGCATCTTGGGTACCGGA	1261
Db		1247	CCATCCCG	AGCTCATGGAGATGGTGAGGACGTGCAGATCGTTCTGCTGGCACCGGCA	1306
QY		1262	AGSCCGCT	ACGAGAAGCTGGTCAAGGCCATCGGCCACCAAGTACAAGGCGCGGCCAAGG	1321
Db		1307	AGAAGA	GTGAGCGCATGCTCATAGCGCCGAGGAGAGTCCACGAGGAGTCCGCG	1366
QY		1322	GCTGGTCA	AGTTCTCGGGCCCTCGGCACATGTCTACCGCCGCGCGGACTTCATCG	1381
Db		1367	CCGTGTG	TCAAGTTCAACCGGCGCTGCGCACCAATCATGGCCGCGCGGCGGCTGCTCG	1426
QY		1382	TGTTGCC	CTCGCGCTTCAGSCCTGCGGCTGATCCAGCTGCACGCCATGCATACGGTA	1441
Db		1427	CCGTACCA	CGCCCTTCGAGCCCTCGGCCCTCATCCAGCTGCAGGGGATTCGATACGGAA	1486
QY		1442	CCGTGCC	CGCTGTAGCTTCCACCGCGGCGTGGTGCACACCGTCAAGAGGGCGTCCACCG	1501
Db		1487	CGCCCTG	CGGCTCGCGGTCCACCGGTGGACTCTCGACACCATCATCGAAGCAACCG	1546
QY		1502	GCTTCCAC	ATGGCGCCCTGA-----ACCCCGACAAGTGTGGACGAGGTGACGCGGACG	1555
Db		1547	GGTTCCAC	ATGGCGCGCTCAGCGTGACTGCAACGTCGTGGAGCGCGGACGCTCAAGA	1606
QY		1556	CCCTGGCG	CCACCGTGGCCGTGCCAGCAGGTGTTTCGGGCGCGCGCTACCCCGAGA	1615
Db		1607	AGGTGGCC	ACCACTTTCAGCGCGCATCAAGTGTGTGCGSACGCCCGGCTACGAGSAGA	1666
QY		1616	TGTTGGCA	AACTGCATCAGCAGGACTGTCTCTGGTCCAAAGCCCGCCCAAGAGTGGAGG	1675
Db		1667	TGTTGAGG	AACTGCATTCAGATCTCTCTCTGGAAGGCGCTGCCAAGAACTGGGAGA	1726
QY		1676	GCCTGCTG	GAGAGGTGGTGTACGGCAAGGGCGGCTGGCCACCGCAAGAGGAGGAGA	1735

RESULT 3

PCT-US02-33122-1

; Sequence 1, Application PC/TUS0233122

GENERAL INFORMATION:

; APPLICANT: ExSeed Genetics

; TITLE OF INVENTION: Starch

FILE REFERENCE:

; CURRENT APPLICATION NUMBER: PCT/US02/331222

; CURRENT FILING DATE: 2002-10-17

; PRIOR APPLICATION NUMBER: 60/329,525

; PRIOR FILING DATE: 2001-10-01

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

```
; LENGTH: 2263
```

```

; TYPE: DNA

```

; ORGANISM: Zea mays

; FEATURE:

OTHER INFORMATION: Wild type sequence EX385

PCT-US02-33122-1

Query Match	22.5%	Score 478;	DB 1;	Length 2263;

Best Local Similarity 59.8%; Pred. No. 1e-56;

QY	161	GTGCCACTTGGCGCGTGACATCGTATGTTGCTGTGAGGTCCGCCCTTGGTGTCCACA	220
Db	408	GCGCCAGCGCGGATGAACCTGCTTCTTCTGCGCGCGAGATGCGCGCTGGAGCAAGA	467
QY	221	CGGCGGCGTGGCGATGTGACTGTGTGCGCTGCTATTGAGTGTGTCGAAGCGCGCCACC	280
Db	468	CCGCGGCGTGGCGAGCTCTCGCGGCGCTTCCGCGGCCATGSCCGCGAACGGGCACC	527
QY	281	CGCTCATGACCAATTGCCCTCGCTACGACCACTAGTCTGAGCGCTTGGACACCTCGGTGG	340
Db	528	GTGTCATGTTGTTCTCTCCCGCTACGACCACTAGTCAAGGACGCGCTTGGACACCGCTCG	587
QY	341	TCGTGGACATCA-----TGGCGGAGAAGTCCGCTACTTCCACTCCATCACA	388
Db	588	TGTCGGAGATCAAGATGGAGACGGGTACGAGACGGTCAGGTTCTTCGACTGCTACAAC	647
QY	389	AGGCGGTGCACCGGTGTGATTGACCACCCCTGGTTCTCGCCAAAGTCTGGGGCAAGA	448
Db	648	GCGGAGTGGACCGGTGTTGTTGACCAACCACACTGTTCTGAGAGGGTTTGGGGAAGA	707
QY	449	CCGCTCCAAAGCTGTACGGCCCCGCTCGCGCGTGTACTCTGACACCAACACAGCGCT	508
Db	708	CCGAGGAGAAGATCTACGGGCTGTGCTTGAACCGGACTACAGGGAACAACCAAGCTCGGT	767
QY	509	TCGCCCTGTTGCAAGCCGCTATTGAGGCTGCCCGGTGCTGCCCTTCGGC-----	561
Db	768	TCAGCCTGCTATGCCAGCGACACTTGAAGTCCAAAGATCCTGAGCCTCAACAACAACC	827
QY	562	-----CCGGCGGAGACTCGCTCTTCGTGGCCAAAGACTGGCACTCCG	604
Db	828	CATACTTCTCGGACCATACGGGAGAGACTGCTGTTGCTGTGCAACGACTTGGACACCG	887
QY	605	CCCTGGTGCCTCTCTGCTGAAGACAGTACCAAGCCCCAAGGGCCAGTTCCACCAAGGCA	664
Db	888	GCCCTCTCTGTTGCTACTCAAGAGCACTACCAGTCCCACGGCATCTACAGGACGCA	947
QY	665	AGTCGGTGTGGCTATCCACAAATCGCTTTCAGGGCCGATGTGGGAGGAGGCTTCA	724
Db	948	AGACCGCTTCTGATCCACAACTCTTACCAAGGCGGTTTCGCTTCTCCGACTACC	1007

QY	845	CTWACAAAGAGATCAACTGGCTGGAAGGGTGGCAATATPCGCCGCCGCAACAGCTGGTGA	904
Db	1080	---GCGGGAGATCAACTGGATGAAGCGCGGGATCCTCAGGCGCAGAGGTCTCTCACCG	1136
QY	905	TGTGCGCCAACTACGGACCGAGATCGCTGCCATGCCCGCGGTGTGGAGCTGGACA	964
Db	1137	TCAGCCCTCTACTAGCGGAGGAGCTCATCTCCGGCATPCGCAAGGGCTCGAGCTCGACA	1196
QY	965	CCGTATCCGCGCAAGGCATTGAGGGCATTTGTGAACGCGCATGGACATTTGAGAGCTGA	1024
Db	1197	ACATCATGCGCTCACCGGCATCACCGGCATCTCAACGCAATGACGTACCGAGTGGG	1256
QY	1025	ACCCAAAGACCGACAAGTTCTCTGTGGCCCTACGACAGCAACAGCGTCTAGCGCGCA	1084
Db	1257	ACCCACGACGGGACAAGTACATCGCGCTGAAGTACGACGTGTGACGCGCGTGGAGGCCA	1316
QY	1085	AGCGCGCGCAAGGAGGCCCTCAGCGCGAGCTGGCGTGTGTGGACCCCAACGCC	1144
Db	1317	AGCGCTGACAAGGAGCGCTCGACGGAGGTCTCGCGGTGGAGCCGGAACATCG	1376
QY	1145	CCCTGTTGCGCTTATCGCGCGCTCGAGGAGCAGAAGGTGTGACATCATCCTGGCCG	1204
Db	1377	CGTGTGTGGCTTATCGCGAGCTGGAAGAGCAGAAGGCCCGCAGCTCATGGCGCG	1436
QY	1205	CCCTGCCCAAGATCTTGGCCACCCCAAG---GTGCAGATCGGCATCTTGGGTACCGCA	1261
Db	1437	CCATCCCGCAGCTCATGGAGATGGTGGAGACGTGAGATCTGTCTGTGGCAGCGGCA	1496
QY	1262	AGCGCGCTACGAGAAGCTGTGCAACGCCATCGCACCAAGTACAAGGCGCGGCCAAG	1321
Db	1497	AGAAGAGTTCGAGCGCATGCTCATGAGCGCCGAGGAGAAGTTCACGAGCAAGTGC	1556
QY	1322	CGTGGTCAAGTTCTCGCGCCCTCGCGCACATGCTCACCGCGCGCGCGCTTCATCG	1381
Db	1557	CCGTGTCAAGTTCAACGGCGCTCGGCGCACCATCATGCGCGCGCGAGTCTCTCG	1616
QY	1382	TGTTGCGCTCGCGCTTCGAGCCCTCGCGCTGATCCAGTGTGACGCCATGCTACGTGA	1441
Db	1617	CCGTACCAACCGCTTCGAGCCCTCGCGCTCATCCAGTTCAGGGAGTCCGATACGGAA	1676
QY	1442	CCGTGCGCTGTGTAGCTTCCACGGCGGCTGTGTGCACACCGTCAAGAGGGCGTCA	1501
Db	1677	CGCCCTCGCGCTCGCGGTCCACGCTGTGACTGTGACACCATCATCGAAGGCAAGCG	1736
QY	1502	GTTTCCACATGGCGCCCTGAACCCG-----ACAAGCTCGACGAGGTGACCGCGAGC	1555
Db	1737	GGTTCCACATGGCGCGCTCAGCGTCACTGTAACTGCTGGAGCGCGGAGCTCAAGA	1796
QY	1556	CCCTGGCGCCACCGTTCGCGCTTCGACGAGGTGTTTGGGGCGGCGCTACCCCGAGA	1615
Db	1797	AGTGTGGCCACCATTTGACGCGGCCATCAAGTGTGCGCACGCGCGGTTCAGAGGA	1856
QY	1616	TGTTGGCCAACTGCATCAGCAGGAGCTGTCTCTGTGTCGAAGCCGCCCAAGTGGGAG	1675
Db	1857	TGTTGAGGAATGCATGATCCAGGATCTCTCTGGAAAGGCCCTGCCAAGACTGGGAGA	1916
QY	1676	GCCTGCTGGAGGAGTGGTGTACGGCAAGGCGGCTGGCCACCGCCAAAGAGGAGGA	1735
Db	1917	ACGTGCTGTCTACGCTCGGGGTGCGC---GGCGGCGAGCCAGGGGTCTGAAGCGAGGA	1973
QY	1736	TCAAGGTGCCGTTCGCGAGAAGATCCCG	1765
Db	1974	TCGCGCGCTCGCCAAAGAGAACGTGCGC	2003

RESULT 5

PCT-US02-33122-2

PC/US02 33122 2
; Sequence 2, Application PC/TUS0233122

GENERAL INFORMATION:

APPLICANT: ExSeed Genetics

; TITLE OF INVENTION: Starch

FILE REFERENCE:

; CURRENT APPLICATION NUMBER: PCT/US02/33122


```
Db      887  -CCGGAAGATCAACTGGATGAAGCGGGATCCTCGAGCGCGACAGGGTCTCTCACCGTC 945
QY      907  TCGCCCAACTAGCGACAGAGATCGTCCCGATGCCCGCGGGTGTGAGCTGGACACC 966
Db      946  AGCCCTTACTAGCGAGAGAGTCTATCTCCGATCGCCAGGGGTGCGAGCTCGACAA 1005
QY      967  GTATCCCGCCCAAGGCAATTGAGGCAATTGAAAGCGCATGGACATGAGGAGTGAAC 1026
Db      1006  ATATCGCGCTCACCGCATACCGGCATCGTCAAGGGATGAGCTGACGAGTGGAC 1065
QY      1027  CCCAGACCGCAAGTTCCTGCTGCGCCCTACGACCAAGAACAGAGCTCTACGCGGCA 1086
Db      1066  CCAGCAGGCAAGATACATCCCGTGAAGTACGAGCTGTACAGCGCGCTGGAGGCAAG 1125
QY      1087  GCGCGCGCAAGAGGCGCTGAGCGGAGCTGGCGCTGCTGTGACCCCGCCCGCC 1146
Db      1126  GCGTGAACAGAGGCGCTGAGCGGAGGCTGCGGCTCCCGGTGAGCGGAATCCCG 1185
QY      1147  CTGTGCGCTTCATCGCGCGCTGGAGGAGCAAGAGGTGTGACATCATCTTGGCGCG 1206
Db      1186  CTGTGTGGCTTCATCGGAGGTGGAAGAGCAAGAGGCGCCCGACGTCTATGCGGCG 1245
QY      1207  CTGCCCCAGATCTGCGCCACCCCAAG---GTGCGAGATCGCCATCTCTGGGTACCG 1263
Db      1246  ATCCCGCAGCTCATGAGATGTGGAGGAGCTGCGAGATCTTCTGTGGGCAAGGCA 1305
QY      1264  GCGCGCTAGAGAAGCTGTGAACGCCATCGGCACCAAGTACAAAGGCGCGCCAAAG 1323
Db      1306  AAGAAGTTCAGCGCATGCTCATGAGCGCGGAGGAGATTCCAGGCAAGTGGCGCG 1365
QY      1324  GTGTGCAAGTTCGCGCGCTGCGGCATGTCACCGCGCGCGCGCGCTACCTGCTG 1383
Db      1366  GTGTGCAAGTTCGAGCGCTGCGGCCTCATCCAGCTGCAGGGGATGCGATACGGA 1425
QY      1384  GTGCGCTCGCGCTGCGCGCTGCGCGCTGATCCAGCTGCAGCGCATGCACTACGTA 1443
Db      1426  GTACACAGCGCTTGCAGCGCTGCGGCCTCATCCAGCTGCAGGGGATGCGATACG 1485
QY      1444  GTGCGCGTGTGAGCTTCCACCGCGCGCTGGTGCAGACCGTCAAGGAGGCGCTCA 1503
Db      1486  CCCTGCGCTGCGGCTCCACCGTGGACTCGTGCACACCATCATCAAGGCAAGACG 1545
QY      1504  TTCACATGGCGCGCTGACCCCG-----ACAGCTGGAGAGGCTGACGCGCGCG 1557
Db      1546  TTCACATGGCGCGCTGACGCTGACGCTGACGCTGACGCTGAGCGCGCGCTCAAG 1605
QY      1558  GTGCGCGCACCGTGGCGGTGCGCGAGGAGTGTGCGGCGCGCGCTACCGGAGATG 1617
Db      1606  GTGCGCACCATTCAGCGCGCATCAAGGTGTGCGGACGCGCGCGCTACGAGGAGT 1665
QY      1618  GTGCGCAACTGCATCAGCAGGACCTGCTCTGTCCAGCGCGCGCGCGCTGAGGAG 1677
Db      1666  GTGAGCAACTGCATCAGGATCTCTCTGAGAGGCGCTCCCAAGAACTGGAGAAC 1725
QY      1678  CTGCTGGAGAGGTGTGTACGGCAAGCGCGGTGGCGCGCGCGCGCGAGAGATC 1737
Db      1726  GTGCTGCTGAGCTCGGGGTCC---GGCGGCGAGCCAGGGGTGCAAGGCGAGGAT 1782
QY      1738  AAGGTGCGCGTGGCGAGAGATCCCGG 1765
Db      1783  GCGCGCTCGCCCAAGGAGAACGTGGCGG 1810
```

RESULT 8

PCN-US02-27129-9

; Sequence 9, Application PC/TUS0227129

; GENERAL INFORMATION:

; APPLICANT: Lananan, Mike

; TITLE OF INVENTION: Self-processing Plants and Plant Parts

; FILE REFERENCE: 109846.317

; CURRENT APPLICATION NUMBER: PCN-US02/27129

; CURRENT FILING DATE: 2002-10-09

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 2223

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic

PCN-US02-27129-9

Query Match 15.6%; Score 331; DB 1; Length 2223;

Best Local Similarity 53.2%; Pred. No. 1.2e-36;

Matches 800; Conservative 0; Mismatches 690; Indels 15; Gaps 4;

```
QY      273  CGSCCAGCGCTCATGACCATTCGCCCTCGCTACGACCACTAGCTGAGCGCTGGACAC 332
Db      717  CGCGCGCAAGGTGTCGACTTCCCTGTACTACAAGATGGACCGCGCTTCGACAACAA 776
QY      333  CTCGTTGCTGTGGACATCATGGCGGAGAGTCCGCTACTTCCATCCCATCAAGAAGG 392
Db      777  GAACATCCCGCGCTCGTGGAGCGCTGAAGAAGCGCGCGCTGCTCCCGCGACCC 836
QY      393  CGTGACCGCGCTGTGATTCACCAACCCCTGTTCTTGGCCCAAGGTCTGGGGCAAGAC 452
Db      837  GTTCAAGGCGGTGACCTTCGTCGCAACCAACGACACGACATCATCTGGAACAAGTAC 896
QY      453  CTCCAAGCTGTACGGCGCGCTCGCGGCTACTACTCTGGACAACCAAGCGCTTCGC 512
Db      897  GCGGTACGCTTCATCTTACCTACGAGGCGGACCCACGATCTTCTACCGGACTACGA 956
QY      513  CCGTTCGCAAGCGCGCTATTGAGGCTGCGCGCTGTGCTGCTTGGCGCGCGCGAGGA 572
Db      957  GGAGTGGCTGAACAAGGACAAGCTCAAGAACCTGATCTGATTCAGGACAACCTCGGG 1016
QY      573  CTCGCTTCTCGTGGCCAAAGACTGCGACTCCGCGCTGGTCCCGCTCTGTGAAGGACA 632
Db      1017  CGGCTCCACTAGTATCTGTACTACGACTCCGAGAGATGATCTTCTCGCGAACGGCTA 1076
QY      633  GTACCAAGCCCAAGGCGGCTTCCAAAGCCAAAGTCTGCTGGCTATCCACAACATCG 692
Db      1077  CGSCTCCAAGCCGCGCTGATCAGGTACATCAACCTGGGCTCTCCAAGGTGGCGCTG 1136
QY      693  CTTCCAGGCGCGCTGTGGAGGAGGCTTTCAGGACACGAAGCTGCCCCAGCGCGCTT 752
Db      1137  GGTGTAGTCCCGAAGTTCGCGCGCGCTGCATCCACGAGTACACCGCGCAACCTCGCG 1196
QY      753  TGACAAGCTGGCGCTTCTCGGAGCGCTATGCCAAGTTCACACTGAGCGCACCCCATG 812
Db      1197  CTGGGTGGACAAGTACTGTACTCTCCGCTGGGTCTACTGAGGCGCGCGCTACGA 1256
QY      813  ---GGAGGACGAGAGCCCGCTGACGGGAAAGACTACAAGAAGATCAACTGGGTGA 869
Db      1257  CCGCGCAACGCGCGCTAGCGTACTCTCGTGTGCTCTACTGCGGCTGGCACATCAT 1316
QY      870  GGTGTCATATCGCGCGCGCAAGCTGTGACTGTGCTGCGCCCACTACGCGACCGAT 929
Db      1317  TGTGTCATCTCTGAGCGCGCACAGGTCTTCAACCGTACGCGCCCTACTACGCGAG 1376
QY      930  CGCTGCGCTGCGCGCGGTGTGGAGCTGACCGCTCATCCGCGCAAGGCGATTTGA 989
Db      1377  CATCTCCGCGATCGCCAGGGGCTGCGAGCTCGACAACATCATGCGGCTTACCGGCA 1436
QY      990  GGCATTTGTGAAGCGCATGACATTGAGGAGTGGAAACCCCAAGACGACAAAGTTCCT 1049
Db      1437  CGGCATCGTCAACGGCATGGACGTCAGCGAGTGGGACCCCGCAGCGGACAAAGTAC 1496
QY      1050  TCGCGCTTACGACCAAGACAGCTCTACCGCCGCAAGGCGCGCGCGCAAGGCGCTG 1109
Db      1497  CGTGAAGTACGAGCTGTGACGCGCGCTGGAGGCGCAAGGCGCTGAACAAGAGGCG 1556
QY      1110  GCGCGAGCTGGCGCTGCTGTGACCGCGCGCGCGCGCGCGCTTCTGCGCTTCTATC 1169
Db      1557  GCGGAGGCTGGCGCTCCCGTGGACCGGACATCCCGCTGGTGGCTTCTATCGCGAG 1616
```


US-10-425-114-17713

Query Match 14.0%; Score 298; DB 8; Length 1406;
Best Local Similarity 59.2%; Pred. No. 4e-32;
Matches 635; Conservative 0; Mismatches 350; Indels 87; Gaps 4;

```
QY 161 GTGCCACTTGGCGCTGGACATCGTGTGCTGAGTGCCTGAGTGCCTTGGTCCCAAGA 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 GCGCCACGCGCGATGAAGTGTCTTCTGCGGCCGAGATGGCGCGTGGAGCAAGA 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 CGSGCGGCTGGCGGATGTGACTGTGGCTTCCCTATTGAGTGGTCAAGCGGGCCACC 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 CGSGCGGCTGGCGGATGTGACTGTGGCTTCCCTATTGAGTGGTCAAGCGGGCCACC 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 GGGTCATGACCATGTCCTGCTAGCAGCAGTACGCTGACGCTGGGACACCTCGGTGG 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 GGTGTCATGTGCTCTCCCGCTAGCAGCAGTACGAGCGCTGGACACCGGTGG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 TCGTGGACATCA-----TGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 TGTCCGAGATCAAGATGGGAGACGGGTAGGACGGTCAAGTCTTCCACTGTCAAGC 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 389 AGSGCGTGACCGGCTGTGATGACACCCCTGTTCTGSCCAAGTCTGGGCAAGA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 GCGGAGTGACCGGCTGTGATGACACCCCTGTTCTGAGAGGGTTTGGGAAAGA 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 449 CGGGCTTCCAGCTGTACGGCCCGCTCGCGGCTGACTACCTGGACACCAAGCGCT 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 CCGAGGAGAAGTCTACGGGCTGTGCTGGACGGCTACAGGACACCGAGTGGGT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 509 TCGCCCTGTCTGCAAGCGGCTATTGAGCTGCCCGCTGCTGCTGCTGCTGCTGCTG 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 TCAGCTGTCTATGCCAGGAGCACTTGAAGCTCCAAAGGATCCTGAGCCTCAACAAC 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 -----CCGCGGAGAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 CATACTTCTCGGACCATACGGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 605 CCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 665 AGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 AGACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 725 AGGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 CGGAGTGAACCTTCCCGGAGAGATTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 785 AGGTTTACACTGAGGCGCACCCCATGAGGAGGAGCAGAGAGCCCGCTGACGGGAAGA 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1046 AG-----CCGCTGGAAG-----1057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 845 CCTAAGAAGATCACTGCTGAAGGTTGGCATATCGCGCCGACAGCTGGTGAAGT 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1058 ---GCGGAGAGATCACTGATGAAGGCGGGATCCTCGAGGCGGACAGGGTCTCAC 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 905 TGTGCGCCAACTACGGACCGGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 TCAGCCCTTACTACGCGGAGGAGTCTATCTCCGGCATCGCAGGGGCTGGAGCTG 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 965 CGGTATCGCGCCAAAGGCGATTGAGGCGATTGTGAACGGCATGGACATGAGAGTGA 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1175 ACATCATGCGCTTCAACCGGCATCACCGGCATCGTCAACGGCATGGAGCTGAGG 1234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1025 ACCCAAGACCGACAGTTCCTGCTGCGCCCTACACAGACACAGCTTACCGCGGCA 1084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1235 ACCCAAGACCGACAGTTCCTGCTGCGCCCTACACAGACACAGCTTACCGCGGCA 1294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1085 AGGCCGCGGCAAGGAGGCGCTCAGGCGGAGCTGGGCTGCTGCTGCTGCTGCTGCT 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1295 AGGCCGCTGAACAAGGAGGCGCTCAGGCGGAGGCTGGGCTCCGCGTGGACCGGAATCC 1354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 1145 CCCTGTTCCGCTTATCGCCGCTGAGGAGCAGAGGCTGTGGACATCAT 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1355 CGCTGTGGGTTTCATCGGAGGCTGAAGAGCAGAGGSCCCCGAGCTCAT 1406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 13

US-10-425-114-16641

; Sequence 16641, Application US/10425114

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 16641

; LENGTH: 1409

; TYPE: DNA

; ORGANISM: zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3066-030-H5_FLI

US-10-425-114-16641

Query Match 14.0%; Score 298; DB 8; Length 1409;

Best Local Similarity 59.2%; Pred. No. 4e-32;

Matches 635; Conservative 0; Mismatches 350; Indels 87; Gaps 4;

```
QY 161 GTGCCACTTGGCGCTGGACATCGTGTGCTGAGTGCCTGAGTGCCTTGGTCCCAAGA 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 GCGCCACGCGCGATGAAGTGTCTTCTGCGGCCGAGATGGCGCGTGGAGCAAGA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 CGSGCGGCTGGCGGATGTGACTGTGGCTTCCCTATTGAGTGGTCAAGCGGGCCACC 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 CGSGCGGCTGGCGGATGTGACTGTGGCTTCCCTATTGAGTGGTCAAGCGGGCCACC 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 GGTGTCATGTGCTCTCCCGCTAGCAGCAGTACGCTGAGCTGGAGCAAGTTCGGTGG 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 GTGTCATGTGCTCTCTCCCGCTAGCAGCAGTACGCTGAGGAGCGGCTGGAGCAAGCTG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 TCGTGGACATCA-----TGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 TGTCGAGATCAAGATGGGAGACGGGTACGAGAGGTCAGGTTCTTCCACTGCTACAGC 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 389 AGSGCGTGACCGGCTGTGATGACACCCCTGCTTCTGAGGCAAGTCTGGGCAAGA 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 GCGAGTGGACCGGCTGTCTGTCACACCCACTGTTCTCTGGAGAGGTTTGGGAAAGA 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 449 CGSGCTCCAAGCTGTAGGCGCCCGCTTCCCGGCGCTGACTACCTGGAGCAAGCAAGGCT 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 689 CCGAGGAGAAGATCTACGGGCTTCTGCTGGAAGGAGTACAGGAGCAAGCAAGCTGCGGT 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 509 TCGCCCTGTCTTCAAGGCGGCTATTGAGGCTGCGCGGCTGCTTCCCTTCCGCTTCCGCT 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 749 TCAGCTGCTATGCCAGGAGCACTTGAAGCTCCAGGATCCTTGAGGCTCAACAACACC 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 -----CCGCGGAGGAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 CATACTTCTCGGACCATACGGGAGGAGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCT 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 605 CCCTGTGCGCGCTCTGCTGAGGACAGTACAGCCCAAGGCGGCTTCCACCAAGGCCA 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 869 GCGCTTCTGCTGCTTACTCAAGAGCAACTACAGTCTCCAGGCGATCTAGAGGAGCCAA 928
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 665 AGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 929 AGACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 2377 TATATGTGTTCTAGACAGATTGACGCTGTGGTCTCATTCAGTTACAAAGCTATCGCTA 2436
QY 1437 CGGTACCGTGGCCGTAGCTCCACCGGGGCTGTGCGACACCGCTCAAGAGGGGT 1496
Db 2437 TGGATCTGACCAATGTTCCTCAACAGGTGATAGTTGACACTGTCAAGAGGGCTT 2496
QY 1497 CACCGGTTCCATATGGGCGGCTGA-----ACCCGACAAAGCTGGAGGCTGACGC 1550
Db 2497 CACTGGAATTCAGATGGGTGCTTCAATGTTGAATGTGATGCTGGATCCGGCTGATG 2556
QY 1551 CGACGCCCTGGCGCCACCGTGGCGGCTGCGCAGCAGGTTGTTGGGGGGCGGCTACCC 1610
Db 2557 GGAATCTATACAAAGACTGTCAAAAGGGCCCTTCAGTCTATGGAACCTCCAGCTTTAC 2616
QY 1611 CGAGATGTGGCAACTGATCAGCAGGACCTGCTGGTCCAAAGCCGCGCAGAGTG 1670
Db 2617 AGAATATCAAGAACTGCATGGCTCAAGATCTTTTCATGGAAGGGCGCTTAAGAGTG 2676
QY 1671 GGAGGCGCTGCTG 1683
Db 2677 GGAGGAAGTCTG 2689

RESULT 15

US-10-336-753-48
; Sequence 48, Application US/10336753
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter L.
; TITLE OF INVENTION: PLANT LIKE STARCHES AND THE METHOD OF MAKING THEM IN
; TITLE OF INVENTION: HOSTS
; FILE REFERENCE: 2461-52
; CURRENT APPLICATION NUMBER: US/10/336,753
; PRIOR FILING DATE: 2003-01-06
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/06660
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/042,939
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-04-04
; NUMBER OF SEQ IDS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 2423
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (Join(1..2094, 2098..2103, 2107..2304, 2308..2421))
US-10-336-753-48

Query Match 10.3%; Score 219.2; DB 9; Length 2423;
Best Local Similarity 50.6%; Pred. No. 2.3e-21;
Matches 801; Conservative 0; Mismatches 673; Indels 108; Gaps 7;
QY 144 TGTACTGTCACATGTCGCACCTGTCGCGCTGACATCGATGATGTTGCTGCTGAGGT 203
Db 588 TGAACCTGGCCCTTTGGCTGGCCCTTAATGTGATGAACGTCGTCGTGGTCTCTGAATG 647
QY 204 CGCCCTTTGGTCCAAAGAGGGGGCGCTGGCGCATGTGACTGTGGCTCGCTATTGAGCT 263
Db 648 TGTCTCTTCTCAAGACAGGTGGCTTGGAGATGTCGTGGGTGCTTTGCCATAGGCTCT 707
QY 264 GCTCAAGCGCGCCACCGGCTCATGACCATTCGCCCTCGCTACGACCATGCTGACGC 323
Db 708 GCGGAGGAGGACACCGGTGTTATGTCGTGATACCAAGATATGAGAGTATGCCAAGC 767
QY 324 CTGGGACACCTCGGTGGCTGTCGACATCATGGCGCAGAGGTCCGCTACTTCCATCTAT 383
Db 768 CCGGG--ATTAGGTGTAAAGAGACGTTACAAAGGTAGCTGGACAGGATTCAGAAGTTACT 825
QY 384 CAAGAAGGCGGTGCACCGCGTGTGGATTGACCAACCCCTGGTTCTCTGGCCAAAGTCTGGG 443

Db 826 TATTTTACTCTTACATTTGATGGAGTTGATTTTGTATCGTAGAAGCCCTCCCTTCCCG 885
QY 444 CAAGACCGGCTCAAGCTGTAGGGCCCGCTCCGGCGCTGACTACTGTGACAAACACAA 503
Db 886 CAC-----CGGCAACAATAATATTATGGGGAGAAAGATTGGATATTGAA 932
QY 504 GGGCTTGGCCCTTTCTGCAAGSCCGCTATTGAGGCTGCCCGGCTGCTGCCCTTCCGCC 563
Db 933 GCGCATGATTTTGTCTGCAAGSCCGCTGTGAGGTTCCATGATGATGCTCCATGTGGCG 992
QY 564 -----CGGCGAGGACTGCTCTTCTGAGGCTGCCCGGCTGCTGCCCTTCCGCC 611
Db 993 TACTGTCTATGATGGCAACTTAGTTTTCATGTTGATGATGATGATGATGATGATGAT 1052
QY 612 GCGCGTCTGCTGAAAGGACGAGTACCAGCCCAAGGCGCAGTTTCAACCAAGGCCAAGTCGGT 671
Db 1053 GCTGTCTATCTAAAGGCTATTAACGGGACATGTTGATGATGATGATGATGATGATGAT 1112
QY 672 GTGGCTATCCACACATCGCTTCCAGGCGCGCATGTGGAGGAGGCTTTCNAAGGACAC 731
Db 1113 GCTGTGTATACACACATTTGCTCATCAGGGTCTGTCGCGCTGTAGACGACTTCGTCAATT 1172
QY 732 GAAGCTGCCCCAGGCGCTTTTCAAGAGCTGGCTTCTCGAGGCGTATGCCAAGGTTTA 791
Db 1173 TGACTTGGCTGAACACTATACGACCACTTCAAACTGTATGATGATGATGATGATGATGAT 1232
QY 792 CACTGAGGCGCCCATGAGGAGGACGAGAGGCGGCGCTGACGGGAAGACCTTACAA 851
Db 1233 C----- 1233
QY 852 GAAGATCAACTGGCTGAAGGTTGGCATTTATCGCCCGCGCACAAAGCTGGTACTGTGCGC 911
Db 1234 ----AGCAACAGTTTTTGTCTCGGGGCTGAAGACGGCAGACCGGCTGTGACCGTTAGCAA 1289
QY 912 CAACTACGCGACCGAGATCGTCCGATGCGCGCGGCTGTGAGCTGGACCTGGACACCGTCA 971
Db 1290 TGGCTACATGTTGGAGCTGAAGACTTCGGAAGCGGCTGGGGCTCCACGACATCAATA 1349
QY 972 CC---CGCCAAAGGGCAITGAGGCAITGTGAACGCAATGGACATGAGAGTGGAAACCC 1028
Db 1350 CCAGAACGACTGGAAGCTGCAGGCGATCGTGAACGCGCATCGACATGAGCGAGTGGAAACCC 1409
QY 1029 CAAGACCGACAAGTTCTGTCGCGCTTACGACCAAGACAGC-----GTCATA 1076
Db 1410 CGCTGTGAGCTGSCACTCCACTCCAGCACTACACCACTACACTTCCAGACGCTGGA 1469
QY 1077 CGCGGCAAGCGCGCGCCCAAGGAGGCGCTGCAAGCGCGAGCTGGGCTGCTGTGACCC 1136
Db 1470 CACCGCAAGCGGCACTGCAAGGCGCGCTGCAAGCGCGAGCTGGGCTGCGAGTCCGCGA 1529
QY 1137 CACCGCGCCCTGTTGCGCTTCAATCGCGCGCTGAGAGGACAGAAAGGTTGGACATCAT 1196
Db 1530 CGACGTGCCACTGATCGGGTTCATCGGGGCTGGACACCAAGAGGGGTGGACATCAT 1589
QY 1197 CTGGCGCCCTGCCAAGATCCTGCGCCACCCCAAGGTGCGAGATGCCATCGCTGGGTAC 1256
Db 1590 CGCGGACGCGATCCACTGGATC---CGGGGAGGAGCGTGCAGCTGATGCTGGGCAC 1646
QY 1257 CGCAAGGCGCGCTACGAGAAGCTGTTGAACGSCATCGGCGCACCAAGTACAAAGGCGCGCG 1316
Db 1647 CGGCGGCGGCACTGAGGACATGCTCGCGCGTTCGAGTCGGACACAGCGACAGGT 1706
QY 1317 CAAGGCGTGGTCAATTTCTCGGCGCGCTGCGGCGACATGCTACCGCGCGGCGCGCACTT 1376
Db 1707 GCGCGGTGGTGGGTTCTCGTGGCGCCCTGCGCGACCGCATCACGCGGCGCGGCGACAT 1766
QY 1377 CATGCTGGTGGCTCGCGCTTCGAGCGCTGCGGCTGATCCAGCTGACGCGCATGACATA 1436
Db 1767 CTGCTGATGCGCGTTCGAGCGCTGCGGGCTGAACAGCTCTACGCGCATGCGGTA 1826
QY 1437 CGGTACCGTGGCGCTGAGCTTCCACCGCGCGCTGGTCGACACCGCTCAAGAGGCGCT 1496
Db 1827 CGGACCGTGGCGCTGTCACCGCGTGGGGGCTCCGCGACACGCTGCGCGCTCGA 1886

QY	1497	CACCGGCTTCCACATGGGGCCCTGACCCCGACAGCTGGACGAGGCTGACGCGGACGC	1556
Db	1887	CCCGTTCAACGACACACCGGCTCGGGTGGACGTTTCGACCGCGGAGCGGAACCGGATGAT	1946
QY	1557	CCTGGCGGCACACGCTGCGCGTGCCACGCGAGGTGTTTGGGGCGCGCTACCCGAGAT	1616
Db	1947	CGACGGCTCTCGCACTGCTCACACAGTACCGGACTACAAGGAGCTGGCGCGCTG	2006
QY	1617	GGTGGCCAACTGCATCAGCCAGGACCTGTCTGTCTCCAAAGCCCGCCAGAAAGTGGAGGG	1676
Db	2007	CAGGGCGCGCGGATGGCGAGGACCTCAGCTGGGACCGCGCGCTGTGTATGAGGA	2066
QY	1677	CCTGCTGGAGGAGTGTGTAC	1698
Db	2067	CGTGTCTGTCAAGGCGAAGTAC	2088

Search completed: June 7, 2003, 02:49:00
Job time : 1012.93 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:49:45 ; Search time 37.5509 Seconds
(without alignments)
2512.365 Million cell updates/sec

Title: US-09-980-771A-3
Perfect score: 3641
Sequence: 1 MAVASTSRPSSARPVINA.....SASKTSAKPLVSAATRKSA 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450	39.8	606	13 AAR25476	Granule-bound star
2	1426	38.2	609	19 AAW56485	Oryza sativa starch
3	1405.5	38.6	563	21 AAG04667	Arabidopsis thalia
4	1405.5	38.6	610	21 AAG04666	Arabidopsis thalia
5	1405.5	38.6	610	23 ABB91229	Herbicidally activ
6	1390	38.2	527	21 AAG04668	Arabidopsis thalia
7	1360.5	37.4	637	19 AAY16604	Protein encoded by
8	1358.5	37.3	533	19 AAW56484	Zea mays waxy gene
9	1341.5	36.8	502	13 AAR25474	Central fragment f
10	1194	32.8	238	22 AAB31175	A granule bound st

11	904	24.8	534	21 AAB49307	Wheat starch synth
12	900.5	24.7	647	20 AAY09004	Wheat starch synth
13	898	24.7	792	23 ABB92160	Herbicidally activ
14	890	24.4	756	21 AAW50818	Wheat soluble star
15	885.5	24.3	649	19 AAW38218	Maize starch synth
16	884.5	24.3	671	19 AAW23937	Wheat soluble star
17	873.5	24.0	626	15 AAR51231	Soluble rice starch
18	872	23.9	652	23 ABB93595	Herbicidally activ
19	869.5	23.9	583	19 AAW56488	Zea mays soluble s
20	865.5	23.8	539	19 AAW56491	Zea mays starch sy
21	860.5	23.6	583	19 AAW70894	Maize starch solub
22	855.5	23.5	804	19 AAW70892	Maize starch solub
23	848	23.3	802	23 AAU99845	Modified barley st
24	848	23.3	813	23 AAU99844	Barley cultivar Mo
25	847.5	23.3	698	19 AAW56487	Zea mays soluble s
26	847	23.3	812	23 AAU99847	Wheat starch synth
27	845	23.2	798	21 AAB37566	Wheat starch synth
28	843	23.2	799	21 AAB37567	Wheat starch synth
29	842	23.2	799	21 AAB37597	Wheat starch synth
30	842	23.1	799	19 AAU23938	Wheat granule-boun
31	826.5	22.7	812	23 AAU99846	Barley line 292 st
32	826	22.7	597	21 AAB37568	Wheat starch synth
33	767	21.1	456	23 AAM51865	Rice starch synth
34	726	19.9	459	17 AAR99540	Soluble starch syn
35	702.5	19.3	669	19 AAW70893	Maize starch solub
36	702.5	19.3	669	19 AAW56486	Zea mays soluble s
37	669	18.4	483	23 ABB54012	Lactococcus lactis
38	647.5	17.8	476	23 ABP27592	Streptococcus poly
39	580	15.9	477	13 AAR25462	giga. Escherichia
40	580	15.9	477	13 AAW53890	E. coli glycogen s
41	570.5	15.7	677	17 AAR99539	Soluble starch syn
42	570.5	15.7	1230	18 AAW17785	Potato tuber solub
43	570.5	15.7	1230	21 AAB49306	Potato starch synt
44	564.5	15.5	1025	23 ABB90967	Herbicidally activ
45	559	15.4	735	18 AAW22729	Z. mays starch syn

ALIGNMENTS

RESULT 1
AAR25476
ID AAR25476 standard; Protein: 606 AA.
XX AAR25476;
AC AC
DT DT
DT 15-JAN-1993 (first entry)
XX DE Granule-bound starch synthase of potato.
DE GBSS gene; amylopectin; amylose production;
KW inhibition; promoter; antisense construct.
XX OS Solanum tuberosum.
XX XX
XX WO9211376-A.
XX PN
XX PD 09-JUL-1992.
XX PF 20-DEC-1991; 91WO-SE00892.
XX PR 21-DEC-1990; 90SE-0004096.
XX PA (AMYL-) AMYLOGENE HB.
XX PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
XX WPI; 1992-250096/30.
XX DR N-PSDB; AAQ26404.
XX PT Modifying potatoes to form amylopectin starch - using an
XX PT anti-sense construct to inhibit granule-bound starch synthase

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

```

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 38.6%; Score 1405.5; DB 21; Length 610;
Best Local Similarity 49.6%; Pred. No. 4.2e-102;
Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;

QY 3 VASTSRPSARPVINAASFVKYKTAQNLRLARGSARKSTSRSAVTGATGATC--ALD 60
DB 49 VDNLQRRSQAPV-----SAKSSKRSKVKTAGKIVCEKGM 85

QY 61 IVMAAEVAPWSKTGLGDLVTGGLPIELVKRGHRVMTIAPRYDQYADADMTSVVVDI-MG 119
DB 86 VIFIGAEVGFWSKTGLGDLVGLGPPALAAARGHRVMTICPRDYQKDAMDTCCVVQIKVG 145

QY 120 EK---VRYFHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHNKRFALFCKA 176
DB 146 DKVENRFFHCYKRGVDRVFDHPFLAKVVGKTSKIYPTGVDYNDNQRLFSLLCOA 205

QY 177 ALEAARVLFF-----GP-GEDCVFVANDWHSALVPLVLLKDEYQPKGQFTKAKSVLAIH 228
DB 206 ALEAPQVLNLSKSYFSGPYGDEVVVANDWHTALLPCYLKSMYQSRGYVMNAKVVFCLH 265

QY 229 NIAFOGRMWEAFKDTKLP---QAAPDKLAFSDGAKVYTEATPMEDEKPLITGKTYKK 285
DB 266 NIAYOGRFAFDYSLNLPISEKSSFD---FMDGYEK-----PVKG---RK 305

QY 286 INWLKGGIITAADKLIVTSFNATETAADAAGGVDELDTVIRAKGIEGVNGMDEEWNPKT 345
DB 306 INWMAALAEARVLTVSPYIAQELISGVDGRGVELHLYLKMKTIVSGIINGMDQVWNPST 365

QY 346 DRFLSAPYQDQSVYAGKAARALQALGELPVDPTAPLFAFICRLEEQKGVDTIILAALPK 405
DB 366 DAYIDIKYDITVTDKAPLKEALQAAVGLPVDROPVPIGFIQRLSEQKGSLLVEAISK 425

QY 406 ILATPKVQJAILGTQKAAEKLVAIGTYKGRAGKVVKFSAPLAHMLTAGADFMVLVPSR 465
DB 426 FMGL-NVQMVILGTGKKMEQAQILEEKEFPKGKAVGAKFNVPVLAHMITAGADFIIVPSR 484

QY 466 FEPGLIQLHMHYGVTPVVAATGGLVDTVKGVTGFHMGALNP--DKLDEADADALAAT 523
DB 485 FEPGLIQLHMHYGVTPVIVASTGGLVDTVKGVTGFHMGALNP--DKLDEADADALAAT 544

QY 524 VRRASEVFAAGRYPEMVANCISQDLSSWSPAQKWEGLLEEVYGVGKGVATAKKEETKVPV 583
DB 545 VTRAVAVYGTSAQEMVNCMPQDSKWKGPALWEKVLVLSLNAV-GSEAGTEGEETI-APL 602

QY 584 AEK 586

```

```

DB 603 AKE 605

RESULT 5
ABB91229
ID ABB91229 standard; Protein; 610 AA.
XX
AC ABB91229;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 440.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 440; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 610 AA;

Query Match 38.6%; Score 1405.5; DB 23; Length 610;
Best Local Similarity 49.6%; Pred. No. 4.2e-102;
Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;

QY 3 VASTSRPSARPVINAASFVKYKTAQNLRLARGSARKSTSRSAVTGATGATC--ALD 60
DB 49 VDNLQRRSQAPV-----SAKSSKRSKVKTAGKIVCEKGM 85

QY 61 IVMAAEVAPWSKTGLGDLVTGGLPIELVKRGHRVMTIAPRYDQYADADMTSVVVDI-MG 119
DB 86 VIFIGAEVGFWSKTGLGDLVGLGPPALAAARGHRVMTICPRDYQKDAMDTCCVVQIKVG 145

QY 120 EK---VRYFHSIKKGVRHWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHNKRFALFCKA 176
DB 146 DKVENRFFHCYKRGVDRVFDHPFLAKVVGKTSKIYPTGVDYNDNQRLFSLLCOA 205

QY 177 ALEAARVLFF-----GP-GEDCVFVANDWHSALVPLVLLKDEYQPKGQFTKAKSVLAIH 228
DB 206 ALEAPQVLNLSKSYFSGPYGDEVVVANDWHTALLPCYLKSMYQSRGYVMNAKVVFCLH 265

QY 229 NIAFOGRMWEAFKDTKLP---QAAPDKLAFSDGAKVYTEATPMEDEKPLITGKTYKK 285
DB 266 NIAYOGRFAFDYSLNLPISEKSSFD---FMDGYEK-----PVKG---RK 305

QY 286 INWLKGGIITAADKLIVTSFNATETAADAAGGVDELDTVIRAKGIEGVNGMDEEWNPKT 345
DB 306 INWMAALAEARVLTVSPYIAQELISGVDGRGVELHLYLKMKTIVSGIINGMDQVWNPST 365

QY 346 DRFLSAPYQDQSVYAGKAARALQALGELPVDPTAPLFAFICRLEEQKGVDTIILAALPK 405
DB 366 DAYIDIKYDITVTDKAPLKEALQAAVGLPVDROPVPIGFIQRLSEQKGSLLVEAISK 425

QY 406 ILATPKVQJAILGTQKAAEKLVAIGTYKGRAGKVVKFSAPLAHMLTAGADFMVLVPSR 465
DB 426 FMGL-NVQMVILGTGKKMEQAQILEEKEFPKGKAVGAKFNVPVLAHMITAGADFIIVPSR 484

QY 466 FEPGLIQLHMHYGVTPVVAATGGLVDTVKGVTGFHMGALNP--DKLDEADADALAAT 523
DB 485 FEPGLIQLHMHYGVTPVIVASTGGLVDTVKGVTGFHMGALNP--DKLDEADADALAAT 544

QY 524 VRRASEVFAAGRYPEMVANCISQDLSSWSPAQKWEGLLEEVYGVGKGVATAKKEETKVPV 583
DB 545 VTRAVAVYGTSAQEMVNCMPQDSKWKGPALWEKVLVLSLNAV-GSEAGTEGEETI-APL 602

QY 584 AEK 586

```

Db	266	NIATQGRFAFDYSLNLNPISEKGSFD---FMQGYEK-----PVKGG---RX	305
QY	286	INWLKGGIIAADKLVTSPNATETAADAAGGVELDTVIRAKGTEGIWNGMDIEWNPXI	345
Db	306	INWKAALIEAHRVLTSPYTAQELISGDRGVELHKYLRKMTVSGIINGMDVQWNEPST	365
QY	346	DKFISAPYDQNSVAGKAAAKEALQAEGLPVDPTAPLFARIGRIEEOGKGVDIILAALPK	405
Db	366	DKYIDIKYDITVTDAKPLINEALQAAVGLPVDKRDVPVIGFIRGLEOQKGSIDLVEAISK	425
QY	406	ILATPKYQIATIGTGKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVPSR	465
Db	426	FMGL-NVOMVLIGTGKKMEAEIILEEKKFPGKAVGAKFNVP LAHMITAGADEFLIIVPSR	484
QY	466	FEPCGLIQLHAMHYGTVPVASTGGLVDTVKGVTFGHMGALNP--DKLDEADADALAAT	523
Db	485	FEPCGLIQLHAMRYGTVPVASTGGLVDTVKDGTYTGFHGRFNKVCVVDPPDDVIATAKA	544
QY	524	VRRASEVFAAGRYPEMVANCISDLSKSPAKQKVEGLEEVVYKGGVATAKKEEIKVPV	583
Db	545	VTRAVAVYGTSAMQEMVKNCKMDQFSNKGPARLWEKVLISLVYA-GSEAGTGEGEI-APL	602
QY	584	AEK 586	
Db	603	AKE 605	
RESULT 6			
AAG04668			
ID	AAG04668 standard; Protein; 527 AA.		
AC	AAG04668;		
DT	17-OCT-2000 (first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 782.		
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
OS	Arabidopsis thaliana.		
PN	EPL033405-A2.		
XX	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-0301439.		
PF	25-FEB-1999; 99US-0121825.		
XX	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	28-APR-1999; 99US-0130891.		
PR	30-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	04-MAY-1999; 99US-0132407.		
PR	05-MAY-1999; 99US-0132484.		
PR	05-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		
PR	06-MAY-1999; 99US-0132487.		
PR	07-MAY-1999; 99US-0132863.		
PR	11-MAY-1999; 99US-0134256.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134221.		
PR	14-MAY-1999; 99US-0134370.		
PR	18-MAY-1999; 99US-01343768.		
PR	19-MAY-1999; 99US-0134941.		
PR	20-MAY-1999; 99US-0135124.		
PR	21-MAY-1999; 99US-0135353.		
PR	24-MAY-1999; 99US-0135629.		
PR	25-MAY-1999; 99US-0136021.		
PR	27-MAY-1999; 99US-0136392.		
PR	28-MAY-1999; 99US-0136782.		
PR	01-JUN-1999; 99US-0137222.		
PR	03-JUN-1999; 99US-0137528.		
PR	04-JUN-1999; 99US-0137502.		
PR	07-JUN-1999; 99US-0137724.		
PR	08-JUN-1999; 99US-0138094.		

```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.

Query Match 38.2%; Score 1390; DB 21; Length 527;
Best Local Similarity 52.7%; Pred. No. 5.6e-101;
Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;

QY 59 LDIVVAAAEVAPWSKTGGLDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSSVVDI- 117
Db 1 MSVFIGAEVGPWSKTGGLDVLGGLPALAARHRVMTICPRYDQYKDAWDTCVVQIK 60
QY 118 MGEK---VRYFHSIKKGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNKKRFALEFC 174
Db 61 VGDKVENVRFFHCYKRGVDRVFDHPIFLAKVVGKTSKIYGPITGVNDNQLRFLSLJC 120
QY 175 KAAIEAARVLPF-----GP-GEDEVFVANVHWSALVFLVLLKDEVPKQFTKAKSVLA 226
Db 121 QAALEAPQVLNLSNKKYSGPYGEDVVFVANVHWTALLPCYLKSMYOSRGVYTNNAKVVEFC 180
QY 227 IHNIAFGQRMWEEAFKDKLP---OAAFDKLFSDGAKYKYVYTEATPMEDEKPPLTGKTY 283
Db 181 IHNIAVQGFADFDDYSLNLPISFKSSFD---FMDGYEK-----PVKG--- 220
QY 284 KKNMLKGGIIAADKLVTVSPNYATEIAADAAGGVELOTVIRAKGTIEGIVNGMDIEWNP 343
Db 221 RKNINWKAILEAHRVLTSPYTAQELISGVDRGVELHKLRYLKMKTYSGLINGMDVQEWNP 280
QY 344 KTDKLSAPYDQNSYAGKAAKEALQAEGLPVDPTAPLFAFVIGRLEQKQVDIILAAL 403
Db 281 STDKYIDIKYDITVTDARPLKEALQAAVGLPVDPRDVPVIGFGRLEQKQSDIILVEAL 340
QY 404 PKILATPKVQIAILGTGKAAEYKLVNAIGTKYKGRAGVYKFSAPLAHMLTAGADPMLVP 463
Db 341 SKFMGL-NYQMVILGTGKKMEQAILEELEKPPGKAVGAKENVPLAHMITAGADIIVP 399
QY 464 SRFEPGGLIQLHAMHYGTVPVVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 521
Db 400 SRFEPGGLIQLHAMRYGTVPIVASTGGLVDTVKGDTYGFHGRFNFKYKCEVWDDDDVIATA 459
QY 522 ATVRRASEVFAGGRYEPYVANCISQDLNWSKPAQKWEGLLEEVYKGGVYATAKKEIKV 581
Db 460 KAVTRAVAVYGTSAQMOEMVKNKMDQDFSWKGPRLWEKVLNLSLVA-GSEAGTEGEI-A 517
QY 582 PVAEK 586
Db 518 PLAKE 522

RESULT 7
AAY16604
ID AAY16604 standard; Protein: 637 AA.
XX
AC AAY16604;
XX
XX 23-AUG-1999 (first entry)
XX
DE Protein encoded by the DNA sequence of the maize waxy gene.
XX
KW Non-glycogen-like polysaccharide production; fermentation; waxy gene;
KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;
KW non-starch branching gene; amylopectin; amylose; plant-like starch.
XX
OS Zea mays.
XX
PN WC09844780-A1.
```


Db 121 SLLCQAALEAPRIILSNPNPFSGPYGEDVVFVVCNDWHITGLPLSKYNSYQSHGIYRDAK 180
QY 223 SVLAHNTAFQGRWEEAFKDTKLPOAAFDKLAFLSDGKAKYVYTEATPMEDEKPLTGKT 282
Db 181 TAFCHNISYQGRFAFSIDYPLNLPFRKSFDFIDGK-----PVEG-- 224
QY 283 YKINLWLGKGLIAADKIVTSPNYATRIADAAGVELDTVIRAKGIEGIVNGMDIEPN 342
Db 225 -RKINWKGILEADRVLTSPYAEELISGARGCELDNIMRLTGITGIVNGMDVSEWD 283
QY 343 PKTDKFLSAPYDQNSVYAGKAAKEALQAEGLPLVDPAPLPAFAGRLBEOQGVDTILAA 402
Db 284 PSRDYIAVKYDVDAEAKALKEALQAEGLVDPDRNIPLVAFIGRLBEOQGVDMAAA 343
QY 403 LPKIL-ATPKVOIALGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADPML 461
Db 344 IPQLMEVVEDVOIIVLLGTGKKFERMLMSAEKFPKRVAVKFNALAHIMAGADVLA 403
QY 462 VPSRFEPCGLIQLHAMHYGTPVYVASTGGLVDTYKGVTFGFMGALNPD--KLDRADADA 519
Db 404 VTSRFEPCGLIQLQGMRYGTPCACASTGGLVDTIIEGKTGFHMGRLSDCNVVEPADVK 463
QY 520 LAATVRASEVFAGRPEWYVANCISQDLSKSPAKWEGLEE--VYVGKGV-----A 572
Db 464 VATTIQRKAVVGTPAYEEMVRNQMIDLSWKGPAKNWENILSLGVAGGEPVGEIEA 523
QY 573 TAKKEIKVP 582
Db 524 PLAKENVAAP 533

RESULT 9
AAR25474
ID AAR25474 standard; Protein; 502 AA.
XX
AC AAR25474;
XX
DT 15-JAN-1993 (first entry)
XX
DE Central fragment from potato GBSS.
XX
KW Granule-bound starch synthase; amylopectin; amylose production;
KW inhibition.
XX
OS Solanum tuberosum.
PN WO9211376-A.
XX
PD 09-JUL-1992.
XX
PF 20-DEC-1991; 91WO-SE00892.
XX
PR 21-DEC-1990; 90SE-0004096.
XX
PA (AMYL-) AMYLOGENE HB.
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstroem O;
XX
DR WPI; 1992-250096/30.
DR N-PSDB; AAQ26401.
XX
PT Modifying potatoes to form amylopectin starch - using an
PT anti-sense construct to inhibit granule-bound starch synthase
XX
PS Claim 1; Page 21; 46pp; English.
XX
CC A genomic library in EMBL3 was prepared using leaves of the potato
CC Bintje. The library was screened with cDNA clones for the 5' and 3'
CC ends of the GBSS gene. A full-length clone of potato GBSS gene was
CC identified (wx311). A BglII-SpeI fragment ("m") of the clone was
CC found to contain the central region of the gene and was cloned in
CC pUC13 to give pSm. Restriction of pSm with NsiI and HpaII gives
CC fragment II which was cloned in pJRD184 to give pJRDmitt. Further

CC restriction of pJRDmitt with HpaII-SstI gives a 2549bp fragment
CC comprising exons and introns from the middle of the gene. The
CC fragment can be used to make antisense constructs to suppress amylose
CC formation in potatoes. See also Q26400-4.
XX
SQ Sequence 502 AA;
Query Match 36.8%; Score 1341.5; DB 13; Length 502;
Best Local Similarity 52.9%; Pred. NO. 3.5e-97;
Matches 275; Conservative 62; Mismatches 146; Indels 37; Gaps 10;
QY 23 GVKKTANQLLRELARGSARKSTSRSAVTGATCATCALDIVMAAEVAPWPKTGGIGDVTG 82
Db 5 GLOSTTN--TKVTPKASRTETKPGCSATIVCGKMNIIFVGTBVGWSKTGGPLDVLG 62
QY 83 GLPIELVKGHRVMTIAPRYDOYADAWDTISVVVDI-MG---EKRYFHSIKKGVRHWID 138
Db 63 GLPPALAAARGHVMITSPRYDOYKDAWDTGVAVEKVGDSIEIFRFHCKYKRGVDRVFD 122
QY 139 HPWFLLAKVWGKTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLPF-----GP-GE 190
Db 123 HPMFLEKVGWKTGSKIYGPKAGLDYLDNLELRELLCQAALAPKVLNLNLSNYSFGPYGE 182
QY 191 DCVFFVANDWHSALVPVLLKDEYQPKQFTKAKSVLAIHNAIFQGRWEEAFKDTKLPOAA 250
Db 183 DVLFTANDWHTALIFCYLKSMYOSRGYIYNATVAFCIHNATFOGRFSDFSDFLLNLPDEF 242
QY 251 FDKLAFSDGYAKVYTEATPMEDEKPPLTGKTKYKINWLGKGLIAADKLVTVSPNYATEI 310
Db 243 RGSDFIDGY-----PVKG---RKINWKGILESHRVTVTVSPYVQEL 283
QY 311 AADAAGVELDTVIRAKGIEGIVNGMDIEWPKYDKFLSAPYD-QNSVYAGKAAKEAL 369
Db 284 VSPVHKGVELDSVLRKTCITGIVNGMDTQWNPATDKYTDYDITTTQYMDAKPLLEAL 343
QY 370 QAEGLPVDPTAPLPAFAGRLBEOQGVDTILAAALPKILATPKVOIALGTGKAAEKLVA 429
Db 344 QAAVGLPVDKTIPLIGFIRLEEQKSDILVAAIHKFIGL-DYQIVVLSTGKKEFEQIE 402
QY 430 AIGTKYKGRAGVVKFSAPLAHMLTAGADFMIVPSRFEPCGLIQLHAMHYGTPVYVASTG 489
Db 403 QLEVLVPNKAAGVAKFNVPLAHMITAGADFMIVPSRFEPCGLIQLHAMRYGTPVICA 452
QY 490 GLVDTVKEGVTGFHMGALNP--DKLDEADADALAAATVRA 527
Db 463 GLVDTVKEGVTGFHMGAFNVECDVDPADVLKIVTIVARA 502

RESULT 10
AAB31175
ID AAB31175 standard; Protein; 238 AA.
XX
AC AAB31175;
XX
DT 02-APR-2001 (first entry)
XX
DE A granule bound starch synthetase II (GBSSII).
XX
KW Granule bound starch synthetase II; GBSSII; starch grain;
KW adenosine diphosphate glucose-alpha1,4-glucan alpha4-glucosyltransferase;
KW starch synthetase.
XX
OS Chlamydomonas reinhardtii.
PN FR2793806-AI.
XX
PD 24-NOV-2000.
XX
PF 21-MAY-1999; 99FR-0006494.
XX
PR 21-MAY-1999; 99FR-0006494.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.

XX PI Dhulst C, Ball S;
XX WPI: 2001-052291/07.
DR N-PSDB; AAC86955.
XX
PT New recombinant nucleic acid encoding fusion of starch synthase and
PT second component, useful in pharmaceutical and food compositions, is
PT targeted to starch granules -
XX
PS Claim 9; Fig 1: 52pp; French.
XX
CC The present sequence represents a granule bound starch synthetase II
CC (GBSSI). The specification describes a recombinant nucleic acid,
CC comprising a polynucleotide encoding an adenosine diphosphate
CC glucose-alpha1,4-glucan alpha4-glucosyltransferase or starch
CC synthetase. Placed upstream of a sequence that encodes a polypeptide
CC of interest. The adenosine diphosphate glucose-alpha1,4-glucan
CC alpha4-glucosyltransferase protein can migrate to sites of
CC biosynthesis of starch grains in plant cells, becoming associated
CC with these grains. The recombinant nucleic acid sequence is used to
CC target polypeptides of interest to starch grains.
XX
SQ Sequence 238 AA;
Query Match 32.8%; Score 1194; DB 22; Length 238;
Best Local Similarity 99.2%; Pred. No. 5.2e-86;
Matches 234; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 473 QLIAMHYGTPVVPVASTGGLVDTVKEGVTGFHMGALNPKLDEADADALAAATVRRASEVFA 532
DQ :|||||
DQ 3 RVHAMHYGTPVVPVASTGGLVDTVKEGVTGFHMGALNPKLDEADADALAAATVRRASEVFA 62
QY 533 GGRYPWVANCISODLSWSKPAOKWEGLLVEVYVYGGVATATKEETKVPVAKIPGDLP 592
DQ :|||||
DQ 63 GGRYPWVANCISODLSWSKPAOKWEGLLVEVYVYGGVATATKEETKVPVAKIPGDLP 122
QY 593 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGFSPAATPKVITYKPAL 652
DQ :|||||
DQ 123 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGFSPAATPKVITYKPAL 182
QY 653 PATAKPTAGLKLAGEASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708
DQ :|||||
DQ 183 PATAKPTAGLKLAGEASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 238
RESULT 11
AAB49307
ID AAB49307 standard; Protein; 534 AA.
XX
XX AAB49307;
XX
XX 01-MAR-2001 (first entry)
XX
XX Wheat starch synthase GBSS protein.
XX
XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
XX food product; adhesive.
XX
XX Triticum aestivum.
XX
XX WO200066745-A1.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000WO-AU00385.
XX
XX 29-APR-1999; 99AU-0000052.
XX
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX (GOOD-) GOODMAN FIELDER LTD.
XX (LINA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX

PI Morell M, Li Z, Rahman S, Appels R;
XX WPI: 2000-647602/62.
XX
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and
XX WST-II, useful in modifying plant starch content and/or composition -
XX Example 15; Fig 9; 21pp; English.
XX
CC The present invention relates to novel protein and coding sequences from
CC wheat. The proteins are wheat starch synthases, designated SSII and
CC SSIII. These can be used in the modification of plant starch content or
CC composition, and to screen plants to identify mutations which affect
CC starch content and composition. The starch can then be used in food
CC products, such as flour, and in films, coatings, adhesives, building
CC materials and packaging materials.
XX
SQ Sequence 534 AA;
Query Match 24.8%; Score 904; DB 21; Length 534;
Best Local Similarity 32.5%; Pred. No. 1.3e-62;
Matches 213; Conservative 48; Mismatches 103; Indels 292; Gaps 9;
QY 64 VAAEVAPWSKTGGGLDVTGGLPIELVKGHRVMTIAPQYADAWDTVVVDIMGEKVR 123
DQ :|||||
DQ 2 VGAEMAPWSKTGGGLDVTGGLPIELVKGHRVMTIAPQYADAWDTVVVDIMGEKVR 49
QY 124 YFHSIKKGVHRVWIDHPWFLAKVWKTGSKLYGPRSGADYLDNKRFAFCFAAEARV 183
DQ :|||||
DQ 50 -----LEKVRGKTKEKIYDPDAGTDYEDNOORFSLCQAALVPR 90
QY 184 L-----PF----- 186
DQ :
DQ 91 LNLNPNYFCHHNISYQGRSFDEFAQLNLPDRFKSSDFIDGYDPVGEKISVYSEI 150
QY 187 -----GP-GEDCVFVANDWHSALVPVLLKDEYOPK 215
DQ :|||||
DQ 151 KVVDKYERVYFHCYKRGVDRVFDHPCFSGPYGDEVVFCNDWHTGLLACYLKSQSN 210
QY 216 GQFTKAKSVLAHNIAFQGRWEEAFKDTLPQAAFDKLAFCSDYAKVYVTEATWEDEK 275
DQ :
DQ 211 GIYRAAKVA----- 219
QY 276 PPLTGTYKKINWLKGGIIAADKLVTSPNYATETAAAGGVELDTVIRAKGIEGVNG 335
DQ :|||||
DQ 220 -----NWKAGILQADKVLTVSPYAYEELISGEARGCELDNIMRLTGITTVNG 268
QY 336 MDIEENPKTKDKFLSAPYDONSIVAGKAAK----- 366
DQ :|||||
DQ 269 MDVSEWDPTKDKFLAVNYDITTALEKALNKLKEEDVQIVLLGTGKKKE 328
QY 367 -----EALQAEGLPVDPTAPLPAFAGRLLEEOKGVDFILA 401
DQ :|||||
DQ 329 RLLKSTEEKFPKVRVAVRENAPLAELQAEVGLPVDKRVPLVAFVIGRLEEOKGVDFILA 388
QY 402 ALPKILLATPKVQIALLGTGKAAVEKLVNAIGTKYKGRAGKGVKFSAPLAHMLTAGDFML 461
DQ :
DQ 389 SIPEI-----HQMAGADVLA 404
QY 462 VPSRFEPCGLIQLHAMHYGTPVVPVASTGGLVDTVKEGVTGFHMGALNPD---KLDEADADA 519
DQ :|||||
DQ 405 VTSRFEPCGLIQLQGMRYGTTPCACSTGGLVDTVIVEGKTGFHMGRLSYDCNVVPEADVK 464
QY 520 LAATVRRASEVFAGGRYPE-----MVANCISODLSWSKPAOKWEGLLVEV 564
DQ :|||||
DQ 465 VVTLKRVKVGTPPAYHEAPLAVENVAAPMVKNCMIODLSWKGPAKNWEDVLEL 520
RESULT 12
AAY09004
ID AAY09004 standard; Protein; 647 AA.
XX
XX AAY09004;
AC

XX 05-JUL-1999 (first entry)
 DT Wheat starch soluble synthase I (SSS I) amino acid sequence.
 DE
 XX Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBSS;
 KW starch branching enzyme; starch soluble synthase; debranching enzyme;
 KW endosperm; wheat; barley; granule-bound synthase; glutenin; starch;
 KW grain softness protein I; bacterial isoamylase; glycogen synthase;
 KW WSB I-D4 gene.
 XX
 OS Triticum tauschii.
 OS WO9914314-A1.
 PN
 XX 25-MAR-1999.
 PD
 XX 11-SEP-1998; 98WO-A000743.
 PF
 XX 20-MAR-1998; 98AU-0002509.
 PR 12-SEP-1997; 97AU-0009108.
 PR
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (GOOD-) GOODMAN FIELDER LTD.
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Li Z, Morell M, Rahman S;
 XX
 XX WPI; 1999-229525/19.
 DR N-PSDB; AAX34651, AAX34652.
 DR
 XX New isolated cereal plant enzyme genes used for, e.g. expression of
 PT antisense sequences of granule bound synthase
 PT
 XX Claim 13; Page 95-97; 171pp; English.
 PS
 XX The invention relates to a novel enzyme of starch biosynthetic pathway
 CC in a cereal plant, where the enzyme is selected from starch branching
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or
 CC SBE I of rice or maize. The methods and products can be used for
 CC targeting expression specifically to the endosperm of the seeds of cereal
 CC plants such as wheat or barley. They can be used for the expression of
 CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low
 CC mol. wt. glutenin, grain softness protein I, bacterial isoamylase,
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They
 CC can be used for modifying the characteristics of starch produced by a
 CC plant. The present sequence represents the wheat SSS I protein sequence.
 XX
 XX Sequence 647 AA;
 Query Match 24.7%; Score 900.5; DB 20; Length 647;
 Best Local Similarity 37.8%; Pred. No. 3.2e-62;
 Matches 225; Conservative 94; Mismatches 205; Indels 71; Gaps 18;
 QY 4 ASTSRPSSARPIVINAAFGVKVTANOLLRELARGSARKSTSRSAVTGATGATCALD--- 60
 DB 78 APAGSPATOPPLPDA---GVGELAPDLLE---GIAEDSIDSIIVAAEQSDSEIMDANE 131
 QY 61 -----IVMAAEVAPWSKGGGLGDTVGGGLPIELVKGRHVTMTAPRY-----DQY 105
 DB 132 QPOAKVTRSFVFTGEAAPYAKSGGLGDTGCSLPIALAARGHVMVMPRYLNGSSDKNY 191
 QY 106 ADAMDTSVVDIM---GEKRYVPHSIKKGVRHWIDHPFLAKVWGKTGSKLYGPRSGA 161
 DB 192 AKALYTGKHIPCFGSGSHVTFPHFYRDNDVWVFDHP-----SYHRPGS-LYGDNFGA 245
 QY 162 DYLDNHRKRALFCKAAETAAKVLDPGP---GEDCVFVANDWHSALVPVLLKDEYQKQGF 218
 DB 246 -FGDNQFRYTLCLYAAACEAPLILELGGYIYGONCMFVYNDWHSALVPVLLAARYKPYGVY 304
 QY 219 TKAKSVLAHNIAPQGRMWEEAFKDTKLPAQAFDKLAFSDGYAKVYITEATPMEDEKPL 278

DB 305 RDSRSTLVITHLAHQGLEPASTYPDLGLPPEWYGALEW-----VPEWARRHALDKG-- 356
 QY 279 TGTYYKKINLWGGIIAADKLVTSNPNYATELAADAAGGVDELDTVI--RAKGIEGIVNOM 336
 DB 357 -----EAVNFKAGVADRTIVTSOGYSWEVIT-AEGGGLNELLSRKSVLNGIVNGI 410
 QY 337 DIEWNPKTDKFLSAPYDQNSVYAGKAAKALQAEPLGPDPTADLFAFIFGRLEQKGY 396
 DB 411 DINDWNPSTDKCLPHHYSVDDI--SGKAKCKAEIQELGLPREDVPLIGFIRGLDYOKGI 469
 QY 397 DIILAALPKILATPKVQIAIILGTGKAAEYKLVNAICTKYGKRAKGVVKTFSAPLAHMLTAG 456
 DB 470 DLIKMAIPELM-REDQFVNLGSGDPIFEGNMHSTESSKDKFRGWGVSVPVSHRITAG 528
 QY 457 ADFMLVPSRFEPCGLIQLHAMHYGTVPVVAAGTGLVDTVK-----EGVTGFHMGAL 507
 DB 529 CDILLMPSRFEPCGLNQLYAMQYGTVPVVGHTGGLRDTVETENPFGAKGEESTGWAFSPL 588
 QY 508 NPKLDEADADALAAIVRRASEVFAGRYPEAVANCISQDLSWSKPAQKWEGLLE 562
 DB 589 TVDKMLWALRTAM-STFREHKPSWEG-----LMKRGMTKDHTWDHAAEQYEQIFE 637

RESULT 13

ABB92160

ID ABB92160 standard; Protein; 792 AA.

XX ABB92160;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1371.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 PS Claim 5; SEQ ID NO 1371; 261pp + Sequence Listing; English.
 CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 SQ Sequence 792 AA;

Query Match 24.7%; Score 898; DB 23; Length 792;

Best Local Similarity 40.5%; Pred. No. 6.9e-62;

342 IRKTIAGIINGMDVQF

```

384 FAFGRLEQKGVDIILAAALPKILATPKVQIAILLGTGKAAYEKLVNLAIGTKYKGRAGVY 443
|||||
402 IGFGRLEQKGSIFVAAALSQ-LVHNQVILILGTGKKFKEQIEHLEVLPDKARGVA 460
|||||
444 KFSAPLAHMITAGADEMLVPSRPEPCGLIQLHAMHYGTVPVASTGGGLVDIVKKGVTGFH 503
|||||
461 KFNYPPLAHMITAGADEMLVPSRPEPCGLIQLHAMRYGTVPVASTGGGLVDIVKKGVTGFQ 520
|||||
504 MGALNP--DKLDEADADALAAVVRASEVFAGGRYPPEMVANCISODLSWSKPAQKWEGLL 561
|||||
521 MGALHVECDKIDSADVAIVKTVARALGTATATAALREMLNCMAQDLSWKGPAPRWKMKL 580
|||||
562 EEVYVGKGVATAKKEIKVPAEKIP 588
|||||
581 LDLEV-TGSEPTGEGEELAPLAKENVP 606
|||||

RESULT 2
YUPOY
starch synthase (EC 2.4.1.21) precursor - potato
N:Alternate names: starch synthase
C:Species: Solanum tuberosum (potato)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
C:Accession: S16555; S24392; S26060; S26061
R:van der Leij, F.R.; Visser, R.G.F.; Ponstein, A.S.; Jacobsen, E.; Feenstra, W.J.
Mol. Gen. Genet. 228, 240-248, 1991
A:Title: Sequence of the structural gene for granule-bound starch synthase of potato (Solanum tuberosum)
A:Reference number: S16555; MUID:91360072; PMID:1886609
A:Accession: S16555
A:Molecule type: DNA
A:Residues: 1-607 <LE1>
A:Cross-references: EMBL:X58453; NID:g21470; PIDN:CAA41359.1; PID:g21471
A>Note: the authors translated the codon AAC for residue 453 as Gly and GCT for residue 454
A:Accession: S24392
A:Molecule type: protein
A:Residues: 78-92, 'X', '94-98, 'XXX', 102, 'XX', 105-107 <LE2>
R:Rohde, W.; Becker, D.; Kull, B.; Salamini, F.
J. Genet. Breed. 44, 311-315, 1990
A:Title: Structural and functional analysis of two waxy gene promoters from potato.
A:Reference number: S26060
A:Accession: S26060
A:Molecule type: DNA
A:Residues: 1-43 <ROH1>
A:Cross-references: EMBL:X52416; NID:g21613; PIDN:CAA36667.1; PID:g21614
A:Experimental source: cv. Granola, clone G1
A:Accession: S26061
A:Molecule type: DNA
A:Residues: 1-43 <ROH2>
A:Cross-references: EMBL:X52417; NID:g21615; PIDN:CAA36668.1; PID:g21616
A:Experimental source: cv. Granola, clone G28
C:Genetics:
A:Gene: waxy
A:Genome: nuclear
A:Introns: 111/3; 138/3; 171/3; 201/3; 223/1; 256/3; 293/2; 374/3; 433/3; 497/3; 526/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing a
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; glycogen/starch biosynthesis; glucosyltransferase; hexosyltransferase
F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNF>
F:78-607/Product: ADPglucose-starch glucosyltransferase #status experimental <MAT>

Query Match 40.2%; Score 1463.5; DB 1; Length 607;
Best Local Similarity 52.7%; Pred. No. 2.2e-79;
Matches 298; Conservative 70; Mismatches 157; Indels 41; Gaps 10;

QY 39 SARKSTSRSAVTCATGATCAGLIDVMAAEVAPSWKTTGGGLDVTGGLPIELVIRGHRVMTI 98
|||
Db 61 ASRTETKPCCSATIVCGKGMNLIIFVTEVGPWSKTGGGLDVLGGLPALAARGHRVMTI 120
|||||
QY 99 AFRYDQYADWNTSVVVDI-MG---EKVYFFHSIKGVHRVWIDHPWFLAKYVGKGTGSKL 154
|||||

```

```

121 SPRYDQYKDWNTSVAVKEVKGVDSTIEIVRFHCYKRGVDRVDFVDPHMFLEKVGWGTGSKI 180
155 YGPRSGADYLDNHRKRFALCKAAEAARVLPF-----GP-GEDCVFVANDHWSALVPV 206
|||||
181 YGPRAGLDYLDNELRFSLLCOALEAPKVLNNSNYSPYGYGVDVLFANDWHTALIPC 240
|||||
207 LLKDEYQPKGQFTKAKSVLAITHNIAFGQRMWEEAFKOTKLPQAAFDKLAISDGYAKVYTE 266
|||||
241 YLKSMYQSRGIYLNKAVAFCHNIAFGQRFSDPFLNLPDEFRGSDFIDGYEK---- 296
|||||
267 ATPMEEDKPLTGKTYKKNWLGIIIAADKLVTSVSNYATEIAADAAGGVELDTVIRA 326
|||||
297 -----PVKG---RKINMKAGILSHRVTVSPYIAQELVSAVDKGVGLDVLRLK 343
|||||
327 KGIEGIVNGMDIEENPKTDKFLSAPYQDQNSVYAGKAAKALQALQELGLPVDPTPLPAF 386
|||||
344 TCITGIVNGMDTOENPATDKYDVKYDITVTMDAKPLLKEALQAAVGLPVDKKTPLIGF 403
|||||
387 IGRLEEQKGVDIILAAALPKILATPKVQIAILLGTGKAAYEKLVNLAIGTKYKGRAGVYKFS 446
|||||
404 IGRLEEQKGSIDLVAALHFKTGL-DVQIVVLGTGKKEFEQETEQLLEVDPNKAAGVAKFN 462
|||||
447 APLAHMITAGADEMLVPSRPEPCGLIQLHAMHYGTVPVASTGGGLVDIVKKGVTGFHMG 506
|||||
463 VPLAHMITAGADEMLVPSRPEPCGLIQLHAMRYGTVPICASTGGGLVDIVKKGVTGFHMG 522
|||||
507 LNP--DKLDEADADALAAVVRASEVFAGGRYPPEMVANCISODLSWSKPAQKWEGLL--- 561
|||||
523 FNVECDVVDPADVLKIVTIVARALAVGYTFLAEAKIKMCMSELSKPEPAKKWETLILGL 582
|||||
562 -----EEVYVGKGVATAKKEIKVP 582
|||||
583 GASGSEPGVGE-EIAPLAKENVATP 607
|||||

RESULT 3
S61504
glycogen(starch) synthase (EC 2.4.1.11) isoform I precursor - garden pea
N:Alternate names: glycogen(starch) synthase
C:Species: Pisum sativum (garden pea)
C:Date: 20-Jul-1996 #sequence_revision 07-Feb-1997 #text_change 16-Aug-2002
C:Accession: S61504; S72372
R:Dry, J.; Smith, A.; Edwards, A.; Bhattacharya, M.; Dunn, P.; Martin, C.
Plant J. 2, 193-202, 1992
A:Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synthase
A:Reference number: S61504; MUID:93251108; PM.D:1302049
A:Accession: S61504
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-603 <DRY>
A:Cross-references: EMBL:X88789; NID:g887570; PIDN:CAA61268.1; PID:g887571
A:Accession: S72372
A:Molecule type: protein
A:Residues: 76-77, 'X', 79, 81-88 <DRW>
C:Superfamily: starch synthase
C:Keywords: glucosyltransferase; hexosyltransferase
F:1-75/Domain: signal sequence #status predicted <SIG>
F:76-603/Product: glycogen(starch) synthase isoform I #status experimental <MAT>

Query Match 39.3%; Score 1431.5; DB 2; Length 603;
Best Local Similarity 50.1%; Pred. No. 1.8e-77;
Matches 291; Conservative 85; Mismatches 158; Indels 47; Gaps 13;

QY 26 KTAQQLRLRELARGSARKSTSRSAVTCATGATCAGLIDVMAAEVAPSWKTTGGGLDVTGGLP 85
|||||
Db 46 RSLNKLHVTRATARGSSDTSEKSLGK--IVCGMSLVFVGAEGVPMWPKTGGGLDVLGGLP 103
|||||
QY 86 IELVKGHRVMTIAPRYDQYADWNTSVVVDI-MGEK---VRYFHSITKGVHRVWIDHPW 141
|||||
Db 104 PVLAGNHRVMTVSPRYDQYADWNTSVVVDI-MGEK---VRYFHSITKGVHRVWIDHPW 163
|||||
QY 142 FLAKVVGKGTGSKLYGPRSGADYLDNHRKRFALCKAAEAARVLPF-----GP-GEDCV 193
|||||

```


Db 164 FLERVWGKTGSKLYGPKTGIDYRDNLRLRFSLICQAALEAPRVLNLSNKKYFSGPYGEDVI 223
 QY 194 EVANDHSAVPVLLKDEYQPKQGTAKSVLAHNTAFQGRMWEAFKD---TKLPQAA 250
 Db 224 FVANDHSAIPCYLKSMSKRSGLYKNAKVAFCIHNTAYQGR---NAFSDFLNLPDEF 280
 QY 251 FDKLAFSDGKLYAKVYTEATPMEDEKPLTKTKKINLWKGIIAADKLVTPSPNYATEI 310
 Db 281 RSSFDYDGNK-----PCGE---KKINWKGIILESQVTVSPHYAKEL 323
 QY 311 ADAAGVVELDVTIRAKGIEGVNGMDIEENPKTDKFLSAPYDONSVAAGKAAKDALQ 370
 Db 324 ISGEDRGVELDNIIRSTGILGVNGMDNREWSPTDRIYDVHNEITVTEAKPLKGLTQ 383
 QY 371 AELGLPVPAPLFAFIRLEEGKGYDIIILALPKILATPKVQIAILGTGKAAYEKLVNA 430
 Db 384 AELGLPVPDSIPILGIFIRLEEGKGSYDILVEALAK-FADENVQIVVLGTGKKIMEKQIEV 442
 QY 431 IGTKYGRKAGVGVKFSAPLAHMLTAGDFMLVPSRFEPCGLIOLHAMHYGTVPVASTGG 490
 Db 443 LEKYPGKALGTFKNSPLAKHIIAGADFIVSRFEPCGLVOLHAMPYGTVPVSVSTGG 502
 QY 491 LVDTVKEGVTGFMGALNP--DKLDEADADALAAATVRRASEVFAGGRYPMVANCISQDL 548
 Db 503 LVDTVKEGVTGFMGALNP--DKLDEADADALAAATVRRASEVFAGGRYPMVANCISQDL 548
 QY 549 SWSKPAQKVEGLLEEV-----VYKGG--VATAKKEEIKVP 582
 Db 563 SWSKPAQKVEGLLEEV-----VYKGG--VATAKKEEIKVP 582

RESULT 4

QJ00703
 N:Altenate names: glycogen(starch) synthase (EC 2.4.1.11) - rice
 C:Species: Oryza sativa (rice)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Aug-2002
 C:Accession: JQ0703
 R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
 Nucleic Acids Res. 18, 5898, 1990
 A:Title: Nucleotide sequence of rice waxy gene.
 A:Reference number: JQ0703; MUID:91016948; PMID:2216792
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <WAX>
 A:Cross-references: EMBL:X53694
 A:Experimental source: strain subsp. japonica Hengfeng
 C:Genetics:
 A:Gene: waxy
 A:Introns: 114/1; 141/1; 174/1; 204/1; 225/2; 259/1; 295/3; 377/1; 436/1; 500/1; 529/1;
 A:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
 C:Superfamily: starch synthase
 C:Keywords: chloroplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltrans

Query Match 39.2%; Score 1427; DB 2; Length 609;
 Best Local Similarity 48.8%; Pred. No. 3.3e-77;
 Matches 301; Conservative 82; Mismatches 172; Indels 62; Gaps 13;
 QY 3 VASTSRPSS-----ARPIVINAASFVGVKKTAN---QLRLARGSAKSTSRSA 48
 Db 18 IADRSAPSLLRHFGQGLKPRSPAGDATSLSVTSARATPKQQRSVQKRSR---FPVS 74
 QY 49 VVGATGATCALDIVMVAEAPVSKTGGGLDVTGGLPIELVKRGRHVMITAPRYQYADA 108
 Db 75 VVIATCA--GMNVFVGAEMAPWSKGGGLDVLGLPLPAPMAANGHRVMVSPRYQYKDA 132
 QY 109 WDTSVVVDI---MGEKRYVFIHSIKGVHRVWIDHPWFLAKVWGKTGSKLYGPRGADYL 164
 Db 133 WDTSVVVAEVKADRYRVERFFHCYKRGVRGDFIDHPSFELEKVMGKTGKIYGPDTGVYK 192
 QY 165 DNHRKRFALFCKAAIEAARVL-----PFGP---GEDCVFVANDHSAIPVLLKDEYQPKG 216

Db 193 DNQRFSLICQAALEAPRVLNLSNKKYFSGPYGEDVI 252
 QY 217 QFTKAKSVLAHNTAFQGRMWEAFKD---TKLPQAA 276
 Db 253 IYRNKAVAFCHNLSYQGRFAFEDYPELNLSEFRSSFDYIDGY-----DTPVEG--- 302
 QY 277 PLTGKTYKKINLWKGIIAADKLVTPSPNYATEIADAAGVVELDVTIRAKGIEGVNGM 336
 Db 303 -----RKINWKGIILESQVTVSPHYAKEL 355
 QY 337 DIEENPKTDKFLSAPYDONSVAAGKAAKDALQ 396
 Db 356 DVSMDPSKDIITAKYDAITAEAKALNEALQAEAGLPVDRKIPLTAFTIGRLEEKGP 415
 QY 397 DIILALPKILATPKVQIAILGTGKAAYEKLVNAIGTKYKGRAGVGVKFSAPLAHMLTAG 456
 Db 416 DVMAAATPELM-QEDVOIVLGTGKKFEKLLKSMEEKYFGKRVAVVKNAPLAHLMAG 474
 QY 457 ADMLVPSRFEPCGLIOLHAMHYGTVPVASTGGIOLHAMHYGTGFMGALNP--DKLE 514
 Db 475 ADVLAVPSRFEPCGLIOLGMRGTACACASTGGIOLHAMHYGTGFMGALNP--DKLE 534
 QY 515 ADADALAAATVRRASEVFAGGRYPMVANCISQDL SWSKPAQKVEGLLEEVVYKGG---- 570
 Db 535 SDVKKVAATILKRAIKVVGTPAYEEMVRNQMQLSWSKPAKMNENVL--LGLGVAGSAPG 592
 QY 571 -----VATAKKEEIKVP 582
 Db 593 IEGDETAPLAKENVAAP 609

RESULT 5

S11481
 N:Altenate names: glycogen(starch) synthase (EC 2.4.1.11) precursor - rice
 C:Species: Oryza sativa (rice)
 C:Date: 05-Mar-1995 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2002
 R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
 Nucleic Acids Res. 18, 5898, 1990
 A:Title: Nucleotide sequence of rice waxy gene.
 A:Reference number: JQ0703; MUID:91016948; PMID:2216792
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <WAX>
 A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA37732.1; PID:g577599
 R:Okagaki, K.J.
 Plant Mol. Biol. 19, 513-516, 1992
 A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.
 A:Reference number: S22519; MUID:92322986; PMID:1377969
 A:Accession: S22519
 A:Molecule type: mRNA
 A:Residues: 1-609 <OKA>
 A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403
 R:Hirano, H.Y.; Sano, Y.
 Plant Cell Physiol. 32, 989-997, 1991
 A:Title: Molecular characterization of the waxy locus of rice (Oryza sativa).
 A:Reference number: S30485
 A:Accession: S30485
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-609 <HIR>
 A:Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401
 A:Accession: PC2190
 A:Molecule type: protein
 A:Residues: 78-113 <HIR>
 A:Experimental source: leaf, cDNA pOSLHC2120
 C:Comment: This protein is involved in amylose synthesis in the rice endosperm.
 C:Genetics:
 A:Gene: waxy
 A:Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/

C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
A;Pathway: starch biosynthesis
C;Superfamily: starch synthase
C;Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>
F:78-609/Product: glycogen (starch) #status experimental <MAT>

Query Match	39.23;	Score 1426;	DB 2:	Length 609;
Best Local Similarity	48.94;	Pred. No. 3.8e-77;		
Matches	302;	Conservative 172;	Indels 62;	Gaps 13;
QY	3	VASTSRPSS-----ARPIVINAAGFVKKTTAN--	OLRLRELARGSARKSTSRSA	48
Db	18	IADRSAPSSLLRHGFOGLKPRSPAGG DATSLSVTTSARATPKQORSVQGRSR---	PFVS	74
QY	49	VTGATGATCALDIYVMAAEVAPWSKTGGGLGVDVTGGLPIELVKRGHRVMTIAPRDQYADA	108	
Db	75	VVYATGA--GMNVYFVGAEMAPWSKTGGGLGVDLGLPLPAMAANGHERVYVISPRIYDQYKDA	132	
QY	109	WDTSVVVDI---NGEKVRYPHSTKKGVHRVWIDHPWFLAKVCKTKGSKLGPGRSGADYL	164	
Db	133	WDTSVVAEIKVADRYERYRPFHCYKRGVDYRFDIHPSEFLKVKWCKTGEKIYGPDTGV DYK	192	
QY	165	DNHKRFALFCKAAIEARV L-----PFGP-----GEDCVFVANDHMSALVPVLLKDEYQPKG	216	
Db	193	DNQMRFSLLCOAALEAPRLNLNNPYPKGTGYGEDVVFVVCVNDWHTGPLASYLKNNYQPNG	252	
QY	217	QFTAKSVLA THNTAFQGRWEEFAFKDKLPQAAFDKLAFSDGAKVYVTEATPMDEDEKP	276	
Db	253	IYRNAKVAFCIHNLSYQGRFAFEDYPELNLSEFRSSEDFIDGY-----DTPVEG-----	302	
QY	277	PLTGKTYKINWLKGGIIAADKLVTVSPNYATEIAADAAGGVELDTVIRAKGIEGVNMG	336	
Db	303	-----RKINWMKAGILEADRVLTVSPYYAEELISGIARGCELDNIMRLGITGVNMG	355	
QY	337	DIEENNPKTDFLSAPYDQNSYAGKAAKAEALQAEGLPVDYPTAPLFAFGRLEEQKV	396	
Db	356	DVSEWSPDKRIYITAKYDATTAIKBAKLNKEALQAEGLPVDYRKTPLTIAFGRLEEQKQP	415	
QY	397	DIIILAAPLKILATPKVQITAILGTGKAAEYKELVNAITGTYKGRAKGVWFSAPLAHMLTAG	456	
Db	416	DVMAAAIPELM--QEDVQVILVLTGCKKFEKLLKSMEEKYPCKVRAVVFKNAPLAHLINAG	474	
QY	457	ADFMLVPSRFPFCGLIQLHAMHYGHVPVASTGGLVDTVKCGVTFGFHMGALNPD--KLDE	514	
Db	475	ADVLAVPSRFPFCGLIQLQGMRYGTTPCASCSTGGVLVDTVIECKTGFHMGRLSVDCVKVYEP	534	
QY	515	ADADALATVRASVFAAGGYPEMVANCI SODLSWSKPAOKWEGLEEVYVKGK-----	570	
Db	535	SDVKVVAATLKRAIKVGTGPAYEVMVRMCNMQDLSWKGPKNWENVL--LGLGVAGSAPG	592	
QY	571	-----VATAKKEEIKVP	582	
Db	593	LEGDEIATPLAKENVAAP	609	

RESULT 6

glycogen(starch) synthase (EC 2.4.1.11) precursor - barley
N;Alternate names: starch synthase
C;Species: Hordeum vulgare (barley)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
C;Accession: S01727; S01728
R;Rohde, W.; Becker, D.; Salamini, F.
Nucleic Acids Res. 16, 7185-7186, 1988
A;Title: Structural analysis of the waxy locus from Hordeum vulgare.
A;Reference number: S01727; MCID:88303345; PMID:2970062
A;Accession: S01727
A;Molecule type: DNA
A;Residues: 1-603 <OH>
A;Cross-references: EMBL:X07931; NID:gl19126; PID:CAA30755.1; PID:q295809
A;Accession: S01728

A;Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-603 <RHOH>
A:Cross-references: EMBL:X07932; NID:g19128; PIDN:CAA30756.1; PID:g19129
C:Genetics:
A:Gene: waxy
A:Introns: 106/3; 133/3; 166/3; 218/1; 251/3; 369/3; 429/3; 493/3; 522/3; 565/3
C:Function:
A:description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra
F;1-72/Domain: transit peptide (amyloplast) #status predicted <NP>
F;73-603/Product: UDPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 39.0%; Score 1421.5; DB 1; Length 603;
Best Local Similarity 51.08; Pred. No. 7e-77;
Matches 291; Conservative 76; Mismatches 167; Indels 37; Gaps 10;

Qy 8 RPSARPIVINAASFGVKKTANOLLRLARGSARKSTSRSVATGATGATCALDVMVAEE 67
Db :
Qy 68 VAPWSTGGLDYDTGGPIELVRKGRHVMTIAPRYDOYADAWTSSVVVDIM---GEKVR 123
Db :
Qy 85 MAPSWTKGGDLGVGLPPAMANGHHVWVSPIRYDOTKDWDTSVLSEIKVADEYERVR 144
Db :
Qy 124 YFHSIKGVHRVTDHPFLAKYWGKTSGLKMPRGADYLNDHKRFALFCAAIEAARY 183
Db :
Qy 145 FFHCYKRGVDRVIDHPWELEKVGKTKKIIGPDAGTDVEDNQRESLICQAALAPRI 204
Db :
Qy 184 L-----PF--GP-GEDCVFYANDHWSALVPVLKDEYOPKGQFTKAKSVLAIHNTAQGR 235
Db :
Qy 205 LNLNNPNFYSGPYGEDVFFVCNDWHMTGLLCYLKSNYSQNGIYYRTAKVFCHINLSIQGR 264
Db :
Qy 236 MWEEAFKDTKLPOAAFDKLFSDGYAKVYVTEATPEEDEPKPTGKTYKKTNWLKGGIIA 295
Db :
Qy 265 FSEDFAQLNLPORFKSSFDFIDGYDK-----PVEG---RKINWMKAGITLQ 307
Db :
Qy 296 ADKLVTSPNVATIADAAGGVVELDVIRAKGTGIVNGMDIEWNPKTDFLSAPYDQ 355
Db :
Qy 308 ADKLVITSYPYAEELISGEARCELDNIMRLTGITGITVINGMDVSEWDPTKDFILAVNYDI 367
Db :
Qy 356 NSVTAGAAAKEALOAEGLPVDPAPTALFAPIGRLEQGKVDDIILAALPKLILATPKVOIA 415
Db :
Qy 368 TTALEAKALKEALOAEVGLPVRKVPVAFIGRLEEQKGDVWMAAIPEILKEEDVQII 427
Db :
Qy 416 ILGTGKAAYEKLWNATIKTKYGRKGVVVKFSAPLAHMLTAGACDMLVPSREPCCGLTLQH 475
Db :
Qy 428 LLGTGKKKFEKLLSMEEKFPKRVVRRFNAPLAHQMMAGADLLAVTSREPECCLTLQ 487
Db :
Qy 476 AMHYGTVPVVASCTGLVDTVKGYTGTFHMGALNPD--KLDEADADALAATAVRRASEFAG 533
Db :
Qy 488 GMRYGTPCVCACSTGGLVDTIIVEGKITGFHMGRSLSDCNVVEPADYKVKVATILKRAVKVGT 547
Db :
Qy 534 GRYPEMVANCISQDSLNSKPAQKWEGGLEEV 564
Db :
Qy 548 PAYOEMKNCKMIODLSWKGPKNWEDVLEL 578
Db :

RESULT 7

starch synthase (EC 2.4.1.21) - sweet potato
N/Alternate names: starch synthase
C/Species: Ipomoea batatas (sweet potato)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C/Accession: Tl0906
R/Wang, S.J.; Yeh, K.W.; Tsai, C.Y.
Submitted to the EMBL Data Library, December 1995
A/Reference number: Z17212
A/Accession: Tl0906
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA

A;Residues: 1-608 <W>
A;Cross-references: EMBL:U44126; NID:g1172158; PID:g1172159
A;Experimental source: cv. Rainong; tuberous root
C;Genetics:
A;Gene: SS67
C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing a
A;Pathway: starch biosynthesis
C;Superfamily: starch synthase
C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 38.9%; Score 1417; DB 2; Length 608;
Best Local Similarity 51.9%; Pred. No. 1.3e-76;
Matches 299; Conservative 62; Mismatches 171; Indels 44; Gaps 11;

QY 32 LRELARGARKTSRSVAATGATGATC---ALDIVMAAEVAPWSKTGGGLDVTGGPIEL 88
Db 52 LRTSARNLAKMEGKMRVQAGTIVCKQOGMNLVFGCEGPKCTGGGLDVLGGLPAL 111
QY 89 VKRGHVRMTIAPRYDOYADAWDTSSVVD-IMG---EKVRYFHSIKKGVRHWIDHPWFLA 144
Db 112 AARGHVRMTVCPRYDOYADAWETCVVVEPVQGVDRIEPVRFFHSYKRGVDRVFDHPMELE 171
QY 145 KWKGTGSKLYGPRSGADYLDNHRKFAALCKAAIEAARVLPF-----GP-CEDCVFVA 196
Db 172 KWKGTGSKLYGPKAGKDYKDNLQRLSLCOAALEAPRVNLNLSKYFSGPYGEDVFFVA 231
QY 197 NDHSHALVPVLKDEYQPKQFTAKSVLAIHNAIFQGRMWEAFKDTKLPOAAFDKLA 256
Db 232 NDHHTALLCYLKTMYQSRGIYMNKAVAFCHNIAYOGRFAFSDFSLNLPDEYKGSFDF 291
QY 257 SDGYAKVYTEANPMEDEKPPITGTYKKINLWLGKGIADKLVTVSPNVAFTIADAAG 316
Db 292 IGIDYK-----PVKG---RKINMKAGIREADRVFTVSPNAYKELVSCVSK 334
QY 317 GVELDTVIRAKGLEGVINGMDIEWNPKDKTSLAPYDONSVAVGAKAAKEALQAEGLP 376
Db 335 GVELDNHIDCGITGICMGMDTQWNPATDKYLVKYDITVYMQAPKLKEALQAAVGLP 394
QY 377 VDPAPLAFAGLEBEQKGVDIILALPKILATPKVQIAILGTGKAAYEKLNAIGTKYK 436
Db 395 VDRNPILGIFGLBEQKGVSDIYLAISFISN-DVQIILGTGKKFQEQEQLVEMVP 453
QY 437 GRAKGVKFSAPLAHMLTAGDFMLVPSRFEPCGLIQLHAMHYGVTPVASTGGGLVDTYK 496
Db 454 DKARGVAKENVPLAHMITAGDFMLPSRFEPCGLIQLHAMRYGTPCICASTGGGLVDTYK 513
QY 497 EGVTFHMGALNPD--KIDEADADALATVRRASEVFAGGRYPPEMVANCISQDLWSKPA 554
Db 514 EGYTFHMGAFNVDCTVPEDVLKVTITVGRALAIYGLAFTEMIKNCSQELSKWGP 573
QY 555 OKWEGILL-----EEVYVGKGVATAKKEIKVP 582
Db 574 KNEITVLLSLVGASEPGVEGE-ETAPLAKENVATP 608

RESULT 8
F86453
granule-bound starch synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F86453
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ausen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F86453
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-610 <STO>
A;Cross-references: GB:AE005172; NID:g6910568; PID:AAF31273.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: starch synthase

Query Match 38.6%; Score 1405.5; DB 2; Length 610;
Best Local Similarity 49.6%; Pred. No. 6.3e-76;
Matches 299; Conservative 76; Mismatches 163; Indels 65; Gaps 14;

QY 3 VASTSRSSARPVINAASFVKTANQLLRELARGARKTSRSVAATGATGATC---ALD 60
Db 49 VDNLQRSQAKPV-----SAKSKRSKSVKTAGKIVCEKGM 85
QY 61 IVMAAEVAPWSKTGGGLDVTGGPIELVKRGRHVRMTIAPRYDOYADAWDTSSVVDI-MG 119
Db 86 VIFGAEGVPSKTGGGLDVLGGLPPALAAARGHVRMTICPRYDOYKDAWDTCVVQIKVG 145
QY 120 EK---VRYFHSIKKGVRHWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKFAALCKA 176
Db 146 DKVENVRFFHCYKRGVDRVFDHPFIPLAKVWGTGSKLYGPRSGADYLDNHRKFAALCKA 205
QY 177 AIEAARVLPF-----GP-CEDCVFVANDHSHALVPVLKDEYQPKQFTAKSVLAIH 228
Db 206 ALEAPQVNLNLSKYFSGPYGEDVFFVANDHHTALLPCYLKSMYQSGVTVNAKVVPCIH 265
QY 229 NIAFGQRMWEAEAFKDKLP---QAAFDKLAESDGYAKVYTEATPMEDEKPPITGTYKK 285
Db 266 NIAYGREFAPDDYSLNLPISFKSSFD---FMDGYEK-----PVKG---RK 305
QY 286 INWLKGLIADKLVTVSPNVAFTIADAAGVVELDTVIRAKGLEGVINGMDIEWNPKT 345
Db 306 INWKAAILEAHVRVTVSPYIAQELISGVDRVELHKYLRMTKTVSGILINGMDYQEWNPST 365
QY 346 DKFLSAPYDONSVAVGAKAAKEALQAEGLPDPAPTAPLFAFAGLEBEQKGVDIILALPK 405
Db 366 DKYIDIKYDITVTDKAPLKEALQAAVGLPVRDVPVIGFVIGLEBEQKGVDIILVEAISK 425
QY 406 ILATPKVQIAILGTGKAAYEKLNAIGTKYKGRAGKGVKFSAPLAHMLTAGDFMLVPSR 465
Db 426 FMGL-NVQMVLITGKTKKMEAOILEEKPFGKAVGVAKENVPLAHMITAGDFMLVPSR 484
QY 466 FEPGGLIQLHAMHYGVTPVASTGGGLVDTYKGVTFHMGALNP--DKLDEADADALAA 523
Db 485 FEPGGLIQLHAMRYGVTPVASTGGGLVDTYKGVTFHMGALNP--DKLDEADADALAA 544
QY 524 VRRASEVFAGGRYPPEMVANCISQDLWSKPAQKWEGLLEEVYVGKGVATAKKEIKVP 583
Db 545 VTRAVAVYGTSAQEMVKNMCDQDFSWKGPALWELKVLNLVA-GSEAGTEGEET-APL 602
QY 584 AEK 586
Db 603 AKE 605

RESULT 9
T14731
N;Altogether starch synthase (EC 2.4.1.11) precursor, granule-bound - sorghum
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C;Accession: T14731
R;Hsing, Y.C.
submitted to the EMBL Data Library, April 1995
A;Reference number: Z18173
A;Accession: T14731
A;Status: preliminary; translated from GB/EMBL/DBDJB
A;Molecule type: mRNA
A;Residues: 1-608 <HS>
A;Cross-references: EMBL:U23945; NID:g1255713; PID:g1255714

A;Experimental source: strain 12311; young seed

C;Genetics:

A;Genome: nuclear

A;Note: wx

C;Function:

A;Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing a

A;Pathway: starch biosynthesis

C;Superfamily: starch synthase

C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltransf

P;1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>

F;78-608/Product: UDPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 37.8%; Score 1376; DB 2; Length 608;

Best Local Similarity 46.7%; Pred. No. 3.6e-74;

Matches 295; Conservative 83; Mismatches 180; Indels 74; Gaps 14;

QY 1 MAVASTSPSSARPIVINAASFGV-----KKTANLLRELARGSARKSTERSAVTGA-- 52

DB 1 MSTLATSQ-----LVATHAGLVGPDSMFRGGVGGLRAARASAAAGDALSMRTSACP 54

QY 53 -----TGATCALTDIVMAAEVAPWSKTGGLGVDVGTGGLPTLVKR 91

DB 55 APROOPAAARRGGRGFRFSLVVCATAGMNVVFGAEMAPWSKTGGLGVDVGLGPPAMAAN 114

QY 92 GHRVMTIAPRYDQYADAWDTSVVDI-MG-----EKVRYFHSIKKKGVHRVWIDHPWFLAKVW 147

DB 115 GHRVWVSPRYDQYKADWDTSVVSEIKMGDGYETVRFHCHYKRGVDRVDFIDPLEFLEWV 174

QY 148 KTGSKLGPSPGADYLDNHKRFALFCKAAATEAARVL-----PF-GP-GEDCVFVANDW 199

DB 175 GKTEKIYGPAGTDYKQNLRFSLCQAALAPRILSLNNPFGSGPYGDDVVFVANDW 234

QY 200 HSALVPVLLKDEYQPKGQFTAKSVLAHNTAFQGRMWEAPKDTKLQPAAPDKLAFSDG 259

DB 235 HTGPLSCYLKSNYQNGIYKDAKTAFCIHNIYSQGRFAFSDDPELNLPERFKSSPFDIDG 294

QY 260 YAKVYTEATPMEDEKPLTGKTYKKNLWKGITIAADKLVTSPNYATEIAADAAGVE 319

DB 295 YEK-----PVGE---RKINMKAGILEADRVLTSPVYAFELISGARGCE 337

QY 320 LDTVIRAKIGIVNGMDIEENKPTDKFLSAPYDONSYYAGKAALQALQELGVPDP 379

DB 338 LDNIMRLTGITIVNGMDVSEWDPSPDKIYAVKYDVSTAVEAKNLKALQAEVGLPVD 397

QY 380 TAPLFAFGRLEEQKQGVDTIILALPKILATPKVQIALLTGKAAEYKLVNAITGTYKGRA 439

DB 398 KIPLVAFGRLEEQKQGVDMVAAIP-LLMEEDIQIVLLGTGKKFERMLMSAEKYPDKV 456

QY 440 KGVVKSAPLAHMLTAGADFMVPSRFPGLIQLHAMHYGTVPVVASTGGLVDVTKGV 499

DB 457 RAVVKFNAALAHIMAGADLLAVTSRFPGLIQLQGMRYGTPFCACASTGGLVDVTIECK 516

QY 500 TGFHMGALNPD--KLDEADADALATVRRASEVAGGRYPVNAVNCISDLSWSKPAOKW 557

DB 517 TGFHMGRLSVDNCNVPEADVKVATLTKRAIKVGTPTAYEEMVKNCMIODLSWKGPKNW 576

QY 558 EGLLEE--VYVYKGG---VATAKKEEIKVP 582

DB 577 ENVLLSLGVAGGEGEGEIEGAIPLAKENVAAP 608

RESULT 10

YUWYU

glycogen(starch) synthase (EC 2.4.1.11) precursor - wheat

N;Alternate names: starch synthase

C;Species: Triticum aestivum (common wheat)

C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002

C;Accession: S16261; S33636

R;Clark, J.R.; Robertson, M.; Ainsworth, C.C.

Plant Mol. Biol. 16, 1099-1101, 1991

A;Title: Nucleotide sequence of a wheat (Triticum aestivum L.) cDNA clone encoding the w

A;Reference number: S16261; MUID:91322506; PMID:1863765

A;Accession: S16261

A;Molecule type: mRNA

A;Residues: 1-615 <CLA>

A;Cross-references: EMBL:X57233; NID:g21901; FIDN:CAA40509.1; PID:g21902

R;Ainsworth, C.; Clark, J.; Baisdon, J.

Plant Mol. Biol. 22, 67-82, 1993

A;Title: Expression, organisation and structure of the genes encoding the waxy protei

A;Reference number: S33636; MUID:93271462; PMID:8499619

A;Accession: S33636

A;Molecule type: protein

A;Residues: 71-75,'X',77-78 <AIN>

C;Genetics:

A;Gene: waxy

C;Function:

A;Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producin

A;Pathway: starch biosynthesis

C;Superfamily: starch synthase

C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra

P;1-70/Domain: transit peptide (amyloplast) #status predicted <TNP>

F;71-615/Product: UDPglucose-starch glucosyltransferase #status experimental <MAT>

Query Match 37.7%; Score 1372; DB 1; Length 615;

Best Local Similarity 49.2%; Pred. No. 6.2e-74;

Matches 279; Conservative 77; Mismatches 159; Indels 52; Gaps 8;

QY 32 LRELARGSARKSTSR-----SNTGATCATCALDIVMAAEVAPWSKTGGLGVDV 82

DB 42 MFTVGASAAPKQSRKPRHFRDRCLSMVVRATCG-GMNLVFGAEMAPWSKTGGLGVDV 100

QY 83 GLPIELVKRGHRVMTIAPRYDOYADAWDTSVVDI-----MGEKVRVYFHSIKKVHRVWID 138

DB 101 GLPMAAANGHRVWISPRYDQYKADWTSVSEIKVVDYRERYRFFCHYKRGVDRVFD 160

QY 139 HPWFLAKVWKGTSKLYGPRSGADYLDNHKRFALFCKAAEAAARVLPF-----GP---- 188

DB 161 HPCFLEKVRGKTKETIYGPDACTDYEDNQRFSLCQAALVPRILDLNNPHFSGPYAM 220

QY 189 -----GEDCVFVANDWHSALVPVLLKDEYQPKGQFTAKSVLAHNTAFQGRMWE 239

DB 221 LCRVPRRAGEDVVFVNCNDWHTGLLACYLKSNSYQNGIYRTAKVAFCIHNIYSQGRSFD 280

QY 240 AFKDTKLQPAAPDKLAFSDGYAKVYTEATPMEDEKPLTGKTYKKNLWKGITIAADKL 299

DB 281 DFAQLNLPRFSSPFDIDGYDK-----PVGE---RKINMKAGILQADKV 323

QY 300 VTVSPNYATEIAADAAGGVVELDTVIRAKIGIVNGMDIEENKPTDKFLSAPYDONS 359

DB 324 LTVSPYAEELISGEARGCELDNIMRLTGITIVNGMDVSEWDPIDKDKELTVNYDVTAL 383

QY 360 AGKAAKALQALQELGVPDTPAPLFAFGRLEEQKQGVDTIILALPKILATPKVQIALLGT 419

DB 384 EGKALNKEALQAEVGLPVDKVPVLAFAFGRLEEQKQGVDMIAAIPVKEEDVQIVLLGT 443

QY 420 GKAAVEKLVNAITGTYKGRAGVWKFSAPLAHMLTAGADFMVPSRFPGLIQLHAMHY 479

DB 444 GKXKFEELLSVEEKFPKTVRAVRENAPLAHQMAGADVLAITSRFEPCGLIQLQGMRY 503

QY 480 GTVPVASTGGLVDVTKGVTFHMGALNPD--KLDEADADALAAVRRASEVAGGRYP 537

DB 504 GTPCACASTGGLVDITVEGTGTFHMGRLSVDNCNVPEADVKVVTTLTKRAVKVVGTPAYH 563

QY 538 EMVANCISDLSWSKPAOKWEGLEEV 584

DB 564 EMVKNCMIODLSWKGPKNWEDVLEL 590

RESULT 11

S07314

glycogen(starch) synthase (EC 2.4.1.11) precursor - maize

N;Alternate names: starch synthase; UDP-glucose starch glycosyltransferase, starch g

C;Species: Zea mays (maize)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Aug-2002

C;Accession: S07314

R;Kloesgen, R.B.; Gierl, A.; Schwarz-Sommer, Z.; Saedler, H.

Mol. Gen. Genet. 203, 237-244, 1986
A:Title: Molecular analysis of the waxy locus of Zea mays.
A:Reference number: S07314
A:Accession: S07314
A:Molecule type: DNA
A:Residues: 1-605 <KLO>
A:Cross-references: EMBL:X03935; NID:g22509; PIDN:CAA27574.1; PID:g1644339
A:Experimental source: line C
A:Note: translation of the nucleotide sequence is not complete
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by sequencing.
C:Genetics:
A:Gene: waxy
A:Introns: 107/3; 134/3; 167/3; 197/3; 219/1; 252/3; 289/2; 370/3; 431/3; 495/3; 524/3;
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
C:Superfamily: starch synthase
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;
F:1-72/Domain: transit peptide (amyloplast) #status predicted <NP>
F:73-605/Product: glycogen (starch) synthase #status experimental <NP>
Query Match 37.6%; Score 1369.5; DB 1; Length 605;
Best Local Similarity 46.7%; Pred. No. 8.6e-74;
Matches 294; Conservative 87; Mismatches 177; Indels 71; Gaps 16;
QY 1 MAVASTSPSPARPPIVINAASFVKKTAANOLLRELARG--SARKSTSRNAVTCATGA--- 55
Db 1 MAALATSQ-----LVATRAGLGV-PDASTFRGAAGGLRGARASAAADTLMSRTSARAA 53
QY 56 -----TCA----LDIVMAAEVAPWSKTTGGLGDDVTGGPLIELVKRGHRV 95
Db 54 PRHQOARRGRFPPLVVCASAGMNVFVGAEMAPWSKTTGGLGDLVGLPLPMAANGHRV 113
QY 96 MTIAPRYDQYADAWTSSVVDI--MG---EKVYFHSIKKGVHVRVTDHPWFLAKVWGKTG 151
Db 114 WYVSPRYDQYKAWDTSSVSEIKMGDGYETVFFHCYKRGVDVDFVDFHLEFLERVWGKTE 173
QY 152 SKLYGPRGADYLDNKHKFPALCKAIEAARVL-----PF--GP-GEDCVFVANDWHSAL 203
Db 174 EKIIYGPVAGTDYRDNQRLFSLLCQALAPRILSLNNNPFYSGPYGDEVVFCVNDWHTGP 233
QY 204 VPVLLKDEYQPGQTTAKSVIAIHNIAFQGRMWEAFKDTKLQPAAFDKLAFSDGYAKV 263
Db 234 LSCYLKSNYQSHGIYRDARTAFICINISYQGRFAFSDFPELNLPERKSSDFIDGYEK- 292
QY 264 YTEATPMEDEXPPDTGTYKIKNLKGLIADKLIVTSPNYATEIADADAGGVLDIV 323
Db 293 -----PVEG---RKINWKGILLEADRVLTIVSPYYABELISGIAIGCELDNI 336
QY 324 IRAKIGIEVINGMDIEENPKTKDFLSAPYDQNSVYAGKAAKALQALGLPDPDTAPL 383
Db 337 MRLTGITGIVNGMDYSEWDPSRDKYIAVYDYSTAVEAKALNKEALQAEVGLPVDNRNPL 396
QY 384 FAFIGRLSEOKGVDIIIALPKIL-ATPKVQTAIIGTGKAAVEKLVNAITGTYKRAKGV 442
Db 397 VAFIGRLSEOKGPDVMAAIPOLMEMVEDVQIVLLGTGKTKKPERMLMSAEKFPCKVRV 456
QY 443 VKFSAPLAHMLTAGADMLVPSRPPCGLIQLHAMHYCTVPVAVTGGGLVDIVKGVTFG 502
Db 457 VKENALAHHLMAGADVLAVTSRFPFCGLIQGMRGYTPCPCASTGGGLVDVPIIEGRTG 516
QY 503 HMGALNPD--KLDEADADALATVRRASEVAGGRYPPEMVANCISODLSWSKPAQKWEGL 560
Db 517 HMGRLSVCNVEPADVKVATTLQRAIKVVGTPAYEEMVRCMQLDLSWKGPAKNWENV 576
QY 561 LEE--VYVKGKV-----ATAKKEIKVP 582
Db 577 LLSLGVAGGEPGVEGEIAPLAKENVAAP 605

RESULT 12

T07921

probable starch synthase (EC 2.4.1.21), granule bound - Chlamydomonas reinhardtii (fragm
N:Alternate names: probable starch synthase

C:Species: Chlamydomonas reinhardtii
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C:Accession: T07921
R:D'hulst, C.; Abel, G.; Kossmann, J.; Ball, S.
submitted to the EMBL Data Library, September 1997
A:Description: Cloning of cDNAs coding for starch synthases in the green algae Chlam.
C:Reference number: Z16218
A:Accession: T07921
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-238 <DHU>
A:Cross-references: EMBL:AF026420; NID:g3169784; PID:g3169785
C:Genetics:
A:Gene: STA2
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producti
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase
Query Match 32.8%; Score 1194; DB 2; Length 238;
Best Local Similarity 99.2%; Pred. No. 6.3e-64;
Matches 234; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 473 QLHAMHYGTVPVVASTGGLVDIVYKEGVTGFHMGALNPKLDEADADALATVRRASEVFA 532
Db 3 RVHAMHYGTVPVVASTGGLVDIVYKEGVTGFHMGALNPKLDEADADALATVRRASEVFA 62
QY 533 GGRYPPEMVANCISODLSWSKPAQKWEGLLEEVYVYGGVATAKKEIKVPVAKIPGDLP 592
Db 63 GGRYPPEMVANCISODLSWSKPAQKWEGLLEEVYVYGGVATAKKEIKVPVAKIPGDLP 122
QY 593 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGANRATTPSGPSAAATPKVTTKPAL 652
Db 123 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGANRATTPSGPSAAATPKVTTKPAL 182
QY 653 PATAKPKTAGLKLAGEASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708
Db 183 PATAKPKTAGLKLAGEASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 238

RESULT 13

T07668

starch synthase (EC 2.4.1.21) SSI precursor - potato

C:Species: Solanum tuberosum (potato)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002

C:Accession: T07668

R:Abel, G.J.W.

submitted to the EMBL Data Library, January 1997

A:Reference number: Z16081

A:Accession: T07668

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-641 <ABE>

A:Cross-references: EMBL:Y10416

A:Experimental source: cv. Desiree; leaf

C:Genetics:

A:Gene: SSI

C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose productin

A:Pathway: starch synthase

C:Superfamily: starch synthase

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 24.8%; Score 903; DB 2; Length 641;

Best Local Similarity 38.1%; Pred. No. 4.8e-46;

Matches 219; Conservative 91; Mismatches 183; Indels 82; Gaps 20;

QY 30 QLLRELARGARKSTSRSAVTG-----ATGATCALDIVMAAEVAPWSKTTGG 76

Db 89 QLIPHSVAGDATWVESHDIIVANDRDLSEDTEEMETPIKLTFTNFIIVTAAAPSKTTGG 148

QY 77 LGDVTGGPLIELVKRGHRVMTIAPRY-----DOYADAWDTSV--VVDMG--EKVRYF 125

149 LGDVGSLPMAAARGHRVMVSPRYLNGFSDSEKYNANAVDLDRATVHCFGDAQVAFY 208
126 HSIKGGVHVRWIDHPFLAKWGKGLYKPRGADYLDNHKRFALFCKAIAERAVLP 185
209 HEYRAGDWFVVDHSSYC-----RPGTP-YGDIYA-FGDQFRFTLLSHAACEAPLVLP 261
186 FGP---GDCVFVANDHSAVLPVLLKDEYQPKQFTKAKSVLAHNTAFQGRMWEAFK 242
262 LGGTYGKCLFLANDHSAVLPVLLKDEYQPKQFTKAKSVLAHNTAFQGRMWEAFK 321
243 DTKLQPAADFCLAFSDGYAKVYATPMEDEKPLTGKTKYKKNLWKGIIAADKLVTV 302
322 NLGLPPQWYGA---EWIPTWARAHLD-----TGCT---VNVLKALAVADRILTV 368
303 SPNTATETAADAAGGVLDTVI--RAKIEGIVNGMDIEENPKTKDKLSAPYDONSUYA 360
369 SQGYSWELTT-PEGGYGLHLLSSRSQSVLNGITNDVNDWNPSTDEHIAHSYNDL-S 426
361 GKAAAKEALQAEGLPVDPTAPLFAFIRGLEEKQGVDIILAAALPKILATPKVQIAILGTG 420
427 GKQCKTDLOKELGLPIRDCPLIGFIRGLDYQKGVDIILSAIPELMQN-DVQVVMGSG 485
421 KAAEKLVAIGTKYKRAKGVYKFSAPLAHMLTAGADFMVPSRFPGLIQLHAMHYG 480
486 EKQYEDMRHTENLTKDFRAWGVFNVPVSHRITAGCDILLMPSPFPCGLNQLYAMRYG 545
481 TYPVASTGGLVDYTK-----EGV---TGFHMGALNPKLDEADALATVRRASV 530
546 TIPIVHSTGGRLDVTYKDNPNYAQEGIGTGTFSPLTSEKL-----LDLTKLAI----- 595
531 FAGGRYPE-----MVANCISQDLSWSKPAQKWE 558
596 ---GYTEHKSWEGLMRGGRDYSWENAAIQYE 627

RESULT 14

S61505
glycogen(starch) synthase (EC 2.4.1.11) isoform II precursor - garden pea
N:Alternate names: glycogen(starch) synthase (EC 2.4.1.11) isoform II precursor; granule
C:Species: Pisum sativum (garden pea)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 16-Aug-2002
R:Accession: S61505; S72373; S72312
R:Dry, I.; Smith, A.; Edwards, A.; Bhattacharyya, M.; Dunn, P.; Martin, C.
Plant J. 2, 193-202, 1992
A:Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synthase
A:Reference number: S61504; MUID:93251108; PMID:1302049
A:Accession: S61505
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-752 <DRY>
A:Cross-references: EMBL:X88790
A:Accession: S72373
A:Molecule type: protein
A:Residues: 58-59, 'H', 61-73 <DRW>
R:Edwards, E.A.
submitted to the EMBL Data Library, June 1995
A:Reference number: S72312
A:Accession: S72312
A:Molecule type: mRNA
A:Residues: 1-85, 'KVALQRELIIQIAERKK', 104-139, 'SSSSGSAVETKRWKHCQQQLC', 160-752 <EDW>
A:Cross-references: EMBL:X88790; MUID:9887572; PID:g9887573
C:Keywords: glycosyltransferase; hexosyltransferase
F:1-57/Domain: signal sequence #status predicted <SIG>
F:58-752/Product: glycogen (starch) synthase isoform II #status experimental <MAT>

Query Match 24.0%; Score 874; DB 2; Length 752;
Best Local Similarity 36.3%; Pred. No. 3.2e-44;
Matches 219; Conservative 89; Mismatches 191; Indels 104; Gaps 17;
QY 3 VASTSPSPSARPIVNAAGFGVKKTAQNLRLARGSAKSTSRSAVTGATGATCALDIV 62
DB 213 ISSITSSLK-----FENFEGANEPSSKEVANAENFESGGKPPPLACTGNVNNII 264

QY 63 MVAEAVPWSKTGGLGVDVTGGLPFIELVKGHFVMTIAPRYQYADAWDTSV--VVDIMGE 120
DB 265 LVSAECAWPKSGTGLGVDVAGSLPKALARRGHEVWLVAPHYGNVAEADHDIGVRKRYKVAQG 324
QY 121 --KVYFHSIKKGVHVRWIDHPFLAKWGKGLYKPRGADYLDNHKRFALFCKAIAER 178
DB 325 DMEVYFFHTYIDGVIVFDSP-----IFRNLESNIY-----GGNRLDILRRVLPCKAAV 375
QY 179 EAARVLP-----FGPEGDCVFVANDHSAVLPVLLKDEYQPKQFTKAKSVLAHNTAFQ 233
DB 376 EWPVHVPCCGICYGOG-NLVFTANDHWHALLFVYLKAYTRDGLMNYTSLVLIENIAHQ 434
QY 234 GRMBEAFKDTKLPOAADFCLAFSDGYAKVYATPMEDEKPLTGKTKYKKNLWKGII 293
DB 435 GRG-----PVEDFNTVLSNGYLDLFRKMDYVGGEH-----FNIPAAGL 473
QY 294 IADKLIVVSPNYATEIAADAAGGVLDVIRAKG--LEGVNGMDIEENPKTKDKLSA 351
DB 474 KTADRIVTVSHGYAMELKT-SEGGWGLHNLINESDWKFRGVNGVDTKDNWFOFDAILYS 532
QY 352 P-----YDQNSVYAGKAAAKEALQAEGLPVDPTAPLFAFIRGLEEKQGVDIILAAALPKIL 407
DB 533 DGYTNVNLKTLTGKROCKAALQRELGLPREDVPLISFIRGLDHQKGVDLIAELIPWMM 592
QY 408 ATPKVOIAILGTGKAAAYEKLVAIGTKYKRAKGVYKFSAPLAHMLTAGADEMLVPSRFE 467
DB 593 -SHDVQLVMLGTGRADLEQMLEKFEFAQHCDC-RSWVGVFSVKMAHRTITAGSDILLMPSRFE 651
QY 468 PGLQLIQLHAMHYGTPVPAVSTGGLVDYTK-----BGVTFGHMGALNPKLDEADALADA 521
DB 652 PCGLNQLYAMSYGTPVPHGVGGLRDTVQFPNPFDESQVGV-----TFDRAEANKLM 703
QY 522 ATVRRASEVFAGRYPEMVANCI-----LWNCLLTYKDYKKSWEGLQERGMQDLSWDNAAQYEEVLVA 747
DB 564 VVY 566
DB 748 AKY 750

RESULT 15

JQ2322
starch synthase (EC 2.4.1.21) precursor - rice
N:Alternate names: starch synthase
C:Species: Oryza sativa (rice)
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 16-Aug-2002
C:Accession: JQ2322; PQ0811
R:Baba, T.; Nishihara, M.; Mizuno, K.; Kawasaki, T.; Shimada, H.; Kobayashi, E.; Ohni
Plant Physiol. 103, 565-573, 1993
A:Title: Identification, cDNA cloning, and gene expression of soluble starch synthase
A:Reference number: JQ2322; MUID:94302151; PMID:7518089
A:Accession: JQ2322
A:Molecule type: mRNA
A:Residues: 1-626 <BAB1>
A:Cross-references: DDBJ:D16202; MUID:g450484; PIDN:BAA03739.1; PID:g450485
A:Accession: PQ0811
A:Molecule type: protein
A:Residues: 114-129 <BAB2>
A:Experimental source: seed
C:Function:
C:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose production
C:Superfamily: starch synthase
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase
F:1-113/Domain: transit peptide (amyloplast) #status predicted <TRS>
F:114-626/Product: ADPglucose-starch glucosyltransferase, 57KD form #status experiment
F:122-626/Product: ADPglucose-starch glucosyltransferase, 55KD form #status experiment
F:440-448/Region: substrate binding #status predicted

Query Match 24.0%; Score 873.5; DB 2; Length 626;
Best Local Similarity 41.5%; Pred. No. 2.6e-44;
Matches 197; Conservative 77; Mismatches 154; Indels 47; Gaps 13;

QY 61 IYVVAEYAPWKTGGLGVDTGGLPIELVYKRGHRVMTIAPRY-----DOYADAWDTSVV 114
Db 135 VVFTGEASPYAKSGGLGVCGSLPTALALRGHRVNVMPRYMNGALNKNFANAFYTEKH 194
QY 115 VDIM---GE-KVRYFHSIKKGVRHWIDHPWFLAKVWGKTGSKLYGPRSCADYLDNHKRF 170
Db 195 IKIPCEGGEHEVTFEHEYRDSVDWVFDHPSY-----HRPGNLYGDNFGA-FGDQFRY 247
QY 171 ALFCRAAIEAARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKGFYKAKSVLAI 227
Db 248 TLLCYAACAPLILELGGYIYGCKMVFVNDWHSALVPVLLAARYPYGVYRDARSVIVI 307
QY 228 HNTAFQGRMWEAFKDTKLPOAFDKLAFSDGYAKYITEATPMEDEKPPPLGKTYKKIN 287
Db 308 HNLAHQGVPEASTYPLGLPPEWYGALEW-----VFPEWARRHALDKG-----EAVN 354
QY 288 WLKGGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RAKGIEGIVNGMDIEWNPKT 345
Db 355 FLKGAVVTADRIVTYSQGSWEVTT-AEGGQGLNELLSSRKSVLNGIVNGIDINDWNPT 413
QY 346 DKFLSAPYQNSVYAGKAAKALQALGLPVDPTAPLFAFGRLEEQKGVDIILAALEPK 405
Db 414 DKELPHYGSVDDL-SGKAKCAELQELGLPIRPDVPPLIGFICRLDYQKGLDIKLAIPD 472
QY 406 ILATPKVQIAILGTGKAAVEKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVLPSP 465
Db 473 LM-RDNIQVFMVLSGDPGEGWMRSTESGYRDKFRGWGFSVPVSHRITAGCDILLMPSP 531
QY 466 FEPCGLIQLHAMHYGTVPVVASTGGGLVDTVK-----EGVTGFHMGALNPKD 511
Db 532 FEPCGLNQLYAMQYGTVPVVGHTGGLRDIVENFNPFAEKGEQGTGWAFFSLTIEK 586

Search completed: June 4, 2003, 14:57:22
Job time : 24.3001 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:50:41 ; Search time 9.50655 Seconds
(without alignments)
3088.950 Million cell updates/sec

Title: US-09-980-771a-3

Perfect score: 3641

Sequence: 1 MAVASTSRPSSARPVINA.....SASKISAAPLVSAATRKSA 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485.5	40.8	608	1	UGST_MANES
2	1466	40.3	608	1	UGST_ANTMA
3	1463.5	40.2	607	1	UGST_SOLTU
4	1431.5	39.3	603	1	UGST_PEA
5	1426	39.2	609	1	UGST_ORYSA
6	1423	39.1	609	1	UGST_ORYGL
7	1421.5	39.0	603	1	UGST_HORVU
8	1417	38.9	608	1	UGST_IPOBA
9	1376	37.8	608	1	UGST_SORBI
10	1372	37.7	615	1	UGST_WHEAT
11	1369.5	37.6	605	1	UGST_WAIZE
12	903	24.8	641	1	UGS2_SOLTU
13	874	24.0	752	1	UGS3_PEA
14	873.5	24.0	626	1	UGS2_ORYSA
15	869	23.9	610	1	UGS2_WHEAT
16	849.5	23.3	788	1	UGS3_SOLTU
17	684.5	18.8	477	1	GLGA_STRPN
18	677.5	18.6	484	1	GLGA_BACST
19	670.5	18.4	485	1	GLGA_LACLA
20	669	18.4	478	1	GLG1_RHIME
21	653.5	17.9	480	1	GLGA_AGRU5
22	652	17.9	480	1	GLGA_THEMA
23	651	17.9	486	1	GLGA_RHTR
24	646.5	17.8	480	1	GLGA_RHTR
25	637	17.5	477	1	GLGA_CLOAB
26	622	17.1	482	1	GLGA_CLOPE
27	622	17.1	484	1	GLGA_VIBCH
28	614.5	16.9	476	1	GLGA_BACHD
29	604.5	16.6	486	1	GLG2_RHIME
30	604	16.6	476	1	GLGA_YERPE
31	602.5	16.5	481	1	GLGA_RHILO
32	587	16.1	477	1	GLGA_ECOLI
33	587	16.1	477	1	GLGA_SALTI

34	580	15.9	477	1	GLGA_SALTY
35	578	15.9	480	1	GLGA_PASMU
36	570.5	15.7	1230	1	UGS4_SOLTU
37	569	15.6	476	1	GLGA_HAEIN
38	560.5	15.4	472	1	GLGA_ANASP
39	555	15.2	461	1	GLGA_FUSNN
40	552.5	15.2	463	1	GLGA_AQUAE
41	536	14.8	465	1	GLGA_SYN7
42	536	14.7	477	1	GLGA_RHOSH
43	536	14.7	492	1	GLG2_ANASP
44	535	14.7	477	1	GLGA_SYNY3
45	527	14.5	444	1	GLGA_DEIRA

ALIGNMENTS

RESULT 1

UGST_MANES

ID	UGST_MANES	STANDARD;	PRT;	608 AA.
AC	Q43784;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Granule-bound glycogen [starch] synthase, chloroplast precursor			
DE	(EC 2.4.1.11).			
GN	WAXY OR GBSS.			
OS	Manihot esculenta (Cassava) (Manioc).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid 1; Malpighiales; Euphorbiaceae; Manihot.			
OX	NCBI_TaxID:3983;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. M.COL.22; TISSUE=tuberous root;			
RX	MEDLINE=94083565; PubMed=8260633;			
RA	Salenuzzaman S.N., Jacobsen E., Visser R.G.F.;			
RT	*Isolation and characterization of a cDNA encoding granule-bound			
RT	starch synthase in cassava (Manihot esculenta Crantz) and its			
RT	antisense expression in potato.*;			
RL	Plant Mol. Biol. 23:947-962(1993).			
CC	!- FUNCTION: RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE IN RESERVE			
CC	STARCH.			
CC	!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =			
CC	UDP + {(1,4)-alpha-D-glucosyl}(N+1).			
CC	!- PATHWAY: Starch biosynthesis.			
CC	!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.			
CC	!- TISSUE SPECIFICITY: SYNTHESIZED IN A NUMBER OF DIFFERENT ORGANS,			
CC	BUT MOST ABUNDANTLY IN TUBERS.			
CC	!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE			
CC	FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR	EMBL: X74160; CAA52273.1;			
DR	InterPro: IPR001296; Glycos_transf_1.			
DR	Fram: PF00534; Glycos_transf_1; 1.			
KW	Glycogen biosynthesis; Transferase; Glycosyltransferase;			
KW	Transit peptide; Chloroplast; Starch biosynthesis.			
FT	TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).			
FT	CHAIN 79 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.			
FT	BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).			
SQ	SEQUENCE 608 AA: 66968 MW: C9C970CD3011BDBD CRC64;			

Query Match 40.8%; Score 1485.5; DB 1; Length 608;
Best Local Similarity 53.6%; Pred. No. 1e-80; Indels 39; Gaps 11;
Matches 304; Conservative 67; Mismatches 157;

```

CC      at the end of the day.
CC      -/- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by, and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to licensee@isb-sib.ch).
CC      -----
CC      EMBL; AJ006293; CAA06958.1; -.
CC      DR EMBL; AJ006294; CAA06959.1; -.
CC      DR InterPro: IPR001296; Glycos_transf_1.
CC      DR Pfam; PF00534; Glycos_Transf1; 1.
CC      KW Glycogen biosynthesis; Transferrase; Glycosyltransferase;
CC      FW transit peptide; Chloroplast; Starch biosynthesis.
CC      FT TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).
CC      FT CHAIN 79 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
CC      FT BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
CC      SQ SEQUENCE 608 AA; 66361 MW; 6527D53D565B6EOC CRC64;
CC
CC      Query Match 40.3%; Score 1466; DB 1; Length 608;
CC      Best Local Similarity 50.1%; Pred. No. 1.5e-79;
CC      Matches 308; Conservative 72; Mismatches 175; Indels 60; Gaps 13;
CC
CC      QY 5 STSRPSARPIVNAASFVK-----K*ANQLRLRLARGAR-----KSTSRSAV 49
CC      DB 17 ATSSPDK-----INLAQVGLRNQOFTNGLRSINMVNDKLMRNNAKQSRSLVKKTNGSP 72
CC      QY 50 TGATGATCALDIYVWAAEAVPWSKTGGLGVDTGGGLPIELVKGRHVMYIAPRYDQYADAW 109
CC      DB 73 LGRILICGTGNLVFLAEVGPWSKTGGLGVDTGGGLPFAMAGNGHRHVMYIAPRYDQYADAW 132
CC      QY 110 DTSVVDI-MG----EKVRYPHSTKKGVRHVMIDHPFLAKVWGTGSKLYPRSGADYLD 165
CC      DB 133 DTSVVVEIKVGDSTETVRFHCYKRGVDYRFVDHPFLFLEKVGWTKSKIYGNAGTYQD 192
CC      QY 166 NHRKRALFCKAAIEARVLDF-----GP-GEDCVFVANDWHSALVPVLLKDEYQPKQG 217
CC      DB 193 NQLRFLSILCOALEAPRVNLNLTSSKYFSGPYGDEVFVANDWHTALLPCYLKSMYQSKGM 252
CC      QY 218 FTRAKSVLAHTNIAFOGRMWEAFKDFKLPQAAFDKLAFGSDGVAKVYTEATPWEDEKPP 277
CC      DB 253 YLHAKVAFCHNTAYOGRFGSSDFCLNLPQPKSFDFDGEK-----P 298
CC      QY 278 LTGKTYKKINLWKGIIADKLVTPSPNYATEIAADAAGGVEDTIVIRKGIIEGVNGMD 337
CC      DB 299 VKG---RKINWKAILESRRVTVSPYYAMELVSGAEKGVLDNVIKTSITGVNGMD 355
CC      QY 338 IEENPKTDKFLSAPYDQNSVYAKAAKALQAEGLPYDPTAPLFAFICRLEEQKQVD 397
CC      DB 356 TQENNPATDKHIDNTDITVMDAKPLLEKALQAAVGLPVKNIPIVGTGFRLEEGKSD 415
CC      QY 398 ILAALPKILATPKVQIAILGTGKAAAYEKVLVNAIGTKYKGRAGVYKFSAPLAHMITAGA 457
CC      DB 416 ILVAAISKVGL-DVQIIILGTCKKFEQIQLEVLVDPKARGVAKFNVPYLAHMITAGA 474
CC      QY 458 DFMVLVSRFPFGGLIQLHAMHYGTVPVVAATGGLVTVTKEGVTGFHMGALNPD--KLDEA 515
CC      DB 475 DFMVLVSRFPFGGLIQLHAMRYGTIPICASTGGLVDTVTGEGTFHMGAFNVCATVDPA 534
CC      QY 516 DADALAAVVRAREVEFAGGRYPBWVANCISQDLSWSKPAOKWGLL-----EEVYVG 567
CC      DB 535 DVQKIATTVERALAAYGSVAYKEMIQNCAQDLSWKGPKNWEKMLLSLGVSGSEPCVDG 594
CC      QY 568 KGVVATAKKEIKVP 582
CC      DB 595 E-EIAPLAKENVATP 608

```

DD

```

218 218  FTKAKSVLAHNTIAFOGRMWEEAFKDTKLQAAFDKLIAFSDGYAKVYTEATPMEDEKPP 277
      YLHAKVAFCIHNTIAYOGRFGSSDFCLLNLDPQFKSSDFDFDGYEK-----P 298
253 253  :||| ||||| ||| ||| |
      :||| ||||| ||| ||| |
278 278  LTKGTYKKINWLKGGIADKLVTSPNYATETAADAAAGGVEDTVIRAKIGIEGVNGMD 337
      :||| ||||| ||| ||| |
299 299  VKG---RKINMWKAGILESDRVVTSPYYAMELVSGAEKGVELDNVIKTSITGIVNGMD 355
      :||| ||||| ||| ||| |
338 338  IEEWPKTDKFLSAPDYQNSVYAKKAAAEKALQAEGLPVDPTAPLFAFTGRLEEOKGVD 397
      :||| ||||| ||| ||| |
356 356  TQWNPATDKDHTNDITTTWDAKPLLKALQAAGLVPDKNIPVTGFTGRLEEOKGS 415
      :||| ||||| ||| ||| |
398 398  IILAAALPKILATPKVQIAILGTGKAAEYKELVNAIGTKYKGRAGVVKFSAPLAHMLTGA 457
      :||| ||||| ||| ||| |
416 416  ILVAAISKFVGL-DVQIILGTGKKFEOQIQELEVLYPDKARGVAKFNVP LAHMLTGA 474
      :||| ||||| ||| ||| |
458 458  DPLMVPSRPEPCGLIQIOLHAMHYGTVPVVAATSGGLVDTVKEGVTGFHMGALNPD--KLDEA 515
      :||| ||||| ||| ||| |
475 475  DPLMVPSRPEPCGLIQIOLHAMRYGTIPICASTGGLVDTVTBEGTFGHMGAFNVECATVDPA 534
      :||| ||||| ||| ||| |
516 516  DADALAATVRASEVTFAGGRYPBMWVANCISQDILSWSKPAQKWGGLL-----EEVYIG 567
      :||| ||||| ||| ||| |
535 535  DVQIKIATTVTERALAAYSVAYKEMIQNCMAQDILSWKGPAPKNWBEKMLLSLGVSGSEPGVDG 594
      :||| ||||| ||| ||| |
568 568  KGVVATAKKEEIKVP 582
      :||| ||| |
595 595  E-EIAPLAKENVATP 608
      :||| ||| |

```

QY	155	YGPRSGADYLHNHKKFALECAAAIEAARVLPF-----GP-GEDCVFVANDHSALVPV	200
Dd	181	YGPKAGLDLNEIRNFRSLQCAALEAPKVLNLNSSNYTSPGYGEDVLFIANDWHTALIPC	240
QY	207	LKDEYQPKGOFTTKAKSVLAITHNTAFORMWEEAFKDTKLPOQAADFCLAFSDGYAKVYTE	266
Dd	241	YLKSMYQSRGIYLNAKVAFCIHTAYOGRFSFSDPLLLNPDEFSGSDFIDGYEK----	296
QY	267	ATPMEEDKPPLTGKTYKKNWLKGIIADKLVTIVSPNTATEIADAAGVELDIVIRA	326
Dd	297	-----PVKG---RKINNMKAGILSHRVVTVSPYYAQELVSADVDKGVGLDSVLRK	343
QY	327	KGIEGIVMGDMTEENWNPKTDKFLSAPYDONSVTYAGKAAAKEALQAEGLGVPDPTAPLEAF	386
Dd	344	TCITGIVNGMDTQEWNPATDKYTVDKYDIITYMDAKPLLKEALQAAVGLPVDKKIFLG	403
QY	387	IGRLEQKVGDIIIALPALPKILLATPKVOIAIIGTGKAAYEKLVINAGIKYKRAKGVVKFS	446
Dd	404	IGRLEQKGSOIILVAAIHKFGL-DVGITVLGTGCKEFQEIEQLGLEVLFPNKAKGAKVEN	462
QY	447	APLAHLMTAGADMFLVPSRFPCGGLIOLHAMHYGTVDPVWASTGGLVDTVKEGVTGFHMA	506
Dd	463	VPLAHMIITAGADMFLVPSRFPCGGLIOLHAMRYGTVPICASTGGLVDTVKEGVTGFHMA	522
QY	507	LNP--DKLDERADADALAATVVRASEVTPAGRGYPDWANCIQSOLSSKSPAOKWEGL---	561
Dd	523	FNVCECDVYDDADVLKIIVTTVARALAVYGTILAFAMKMKNCSELSKEPAKKWETLLGL	582
QY	562	----EENVYKGKGVATAKKEIKVP	582
Dd	583	GASGSEPGVEGE-ETAPLAKENVATP	607

RESULT 4

ID	UGST_PEA	STANDARD;	PRT;	603 AA.
AC	Q43092;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Granule-bound glycogen [starch] synthase, chloroplast precursor			
DE	(EC 2.4.1.11) (GBSSI).			
OC	Pisum sativum (Garden pea).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OX	eurosid I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.			
OX	NCBI_TaxId=3888;			
RP	[1]			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 76-88.			
RC	STRAIN=cv. BCI/RR; TISSUE=Embryo;			
RX	MEDLINE=93251108; PubMed=1302049;			
RT	Dry I., Smith A., Edwards A., Bhattacharya B., Dunn P., Martin C.;			
RT	"Characterization of cDNAs encoding two isoforms of granule-bound			
RT	starch synthase which show differential expression in developing			
RT	storage organs of pea and potato."			
RL	Plant J. 2:193-202(1992).			
CC	-!- FUNCTION: MAY BE RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE.			
CC	-!- CATALYTIC ACTIVITY: UDP-glucose + {(1.4)-alpha-D-glucosyl}(N) =			
CC	UDP + {(1.4)-alpha-D-glucosyl}(N+1).			
CC	-!- PATHWAY: Starch biosynthesis.			
CC	-!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND,			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN PODS AND LEAVES. NO EXPRESSION			
CC	IN FLOWERS OR STIPULES.			
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF EMBRYONIC			
CC	DEVELOPMENT WITH HIGHEST LEVELS IN LATER DEVELOPMENTAL STAGES.			
CC	-!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCAGEN SYNTHASE			
CC	FAMILY			


```

Query Match      39.2%; Score 1426; DB 1; Length 609;
Best Local Similarity 48.9%; Pred. No. 3.5e-77;
Matches 302; Conservative 81; Mismatches 172; Indels 62; Gaps 13;

QY 3 VASTSRPSS-----ARPIVINAASFVKKKTAN---QLLRELARGSARKSTSSEA 48
DB 18 IADRSAPSSLRHFGQGLKPRSPAGGDATLSVTTTSARATPKQORSVQGRSRR---FPVS 74

QY 49 VTGATGATCADIWVAEVAEPWSKTGGGLDVTGGLPIELVKGHRVMTIAPRYDQYADA 108
DB 75 VVIATGA--GMNVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVWVISPRIYDQYKDA 132

QY 109 WDTSVVVDI-----MGEKVRYFHSIKKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYL 164
DB 133 WDTSVVAEIKVADRYERVRFFHCYKRGVDRVIDHPSELEKVGWGTGKIYGPDTGVQYK 192

QY 165 DNHRKRALFCKAAIEAARVL-----PFGP---GEDCVFVANDWHSALVPLLLKDYQPKG 216
DB 193 DNQMFSLLCQAALAPILNINNPNYFKGTGDEGVFVNDWHTGGLPLASYLKNNYQPNG 252

QY 217 QFTKAKSVLAIHNIAPQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYTEATPMEDEKP 276
DB 253 IYRNAKVAFCIHNISYQGRFAFEDYPENLSERFSRSDFDIDY-----DTPVEG---- 302

QY 277 PLTGKTYKINWLKGGITAAADKLIVTVSPNYATEIAADAAGGVDELDTVIRAKGIEGIVNGM 336
DB 303 -----RKINWMKAGILEADRLVTVSPYIAELISGIARGCELDNIMLTGITGIVNGM 355

QY 337 DIEENPKTKFLSAPYDQNSYVAGKAAKALQAEGLPVDPTAPLFAFIRGLREEQKV 396
DB 356 DVEWDPSKDYITAKYDATTATIAEAKLNKEALQAEGLPVDKRIPLTAFIRGLREEQKV 415

QY 397 DIILAALPKILATPKVQITAILGTGKAAYEKLVNAIGTKYKRAKGVKFSAPLAHLTAG 456
DB 416 DVMAAAIPELM--QEDVQIVLLGTGKKFELKLSMEEKYPGKVRVAVVFNAPLAHLIMAG 474

QY 457 ADEMLVPSRFPCCGLIQLHAMHYGTVPVVASTGGGLVDTVIEGTFHMGRLSVDCVKVVEP 514
DB 475 ADVLAVPSRFPCCGLIQLQGRYGTIPCACTGGGLVDTVIEGTFHMGRLSVDCVKVVEP 534

QY 515 ADADALAATVRRASEVFAAGGRYPENVANCISODLSWSKPAQKWEGLLEEVYVYKGG---- 570
DB 535 SDVKKVAATLKRAIKVGTGPAYEEMVRNOMDLSWKGPKNWENVL--LGLGVAGSAPG 592

QY 571 -----VATAKKEIKVP 582
DB 593 IEGDEIAPLAKENVAAP 609

RESULT 6
UGST_ORYGL STANDARD; PRT: 609 AA.
AC Q42968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY.
OS Oryza glaberrima (African rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4538;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. GMS1;
RX MEDLINE=92134825; PubMed=1685658;
RA Umeda M., Ohtsubo H., Ohtsubo E.;
RT "Diversification of the rice Waxy gene by insertion of mobile DNA
RL Jpn. J. Genet. 66:569-586(1991).
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.

```

```

-!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
UDP + [(1,4)-alpha-D-glucosyl](N+1).
-!- PATHWAY: Starch biosynthesis.
-!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
-!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL: D10472; BAA01272.1;
DR InterPro: IPR001296; Glycos_transf_1;
DR Pfam: PF00534; Glycos_transf_1;
KW Glycogen biosynthesis; transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 77
FT CHAIN 78 609
FT BINDING 97 97
FT SEQUENCE 609 AA; 66475 MW; C228BFB9C407FA5 CRC64;
DR
Query Match 39.1%; Score 1423; DB 1; Length 609;
Best Local Similarity 48.9%; Pred. No. 5.2e-77;
Matches 302; Conservative 80; Mismatches 173; Indels 62; Gaps 13;

QY 3 VASTSRPSS-----ARPIVINAASFVKKKTAN---QLLRELARGSARKSTSSEA 48
DB 18 IADRSAPSSLRHFGQGLKPRSPAGGDATLSVTTTSARATPKQORSVQGRSRR---FPVS 74

QY 49 VTGATGATCADIWVAEVAEPWSKTGGGLDVTGGLPIELVKGHRVMTIAPRYDQYADA 108
DB 75 VVIATGA--GMNVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVWVISPRIYDQYKDA 132

QY 109 WDTSVVVDI-----MGEKVRYFHSIKKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYL 164
DB 133 WDTSVVAEIKVADRYERVRFFHCYKRGVDRVIDHPSELEKVGWGTGKIYGPDTGVQYK 192

QY 165 DNHRKRALFCKAAIEAARVL-----PFGP---GEDCVFVANDWHSALVPLLLKDYQPKG 216
DB 193 DNQMFSLLCQAALAPILNINNPNYFKGTGDEGVFVNDWHTGGLPLASYLKNNYQPNG 252

QY 217 QFTKAKSVLAIHNIAPQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYTEATPMEDEKP 276
DB 253 IYRNAKVAFCIHNISYQGRFAFEDYPENLSERFSRSDFDIDY-----DTPVEG---- 302

QY 277 PLTGKTYKINWLKGGITAAADKLIVTVSPNYATEIAADAAGGVDELDTVIRAKGIEGIVNGM 336
DB 303 -----RKINWMKAGILEADRLVTVSPYIAELISGIARGCELDNIMLTGITGIVNGM 355

QY 337 DIEENPKTKFLSAPYDQNSYVAGKAAKALQAEGLPVDPTAPLFAFIRGLREEQKV 396
DB 356 DVEWDPSKDYITAKYDATTATIAEAKLNKEALQAEGLPVDKRIPLTAFIRGLREEQKV 415

QY 397 DIILAALPKILATPKVQITAILGTGKAAYEKLVNAIGTKYKRAKGVKFSAPLAHLTAG 456
DB 416 DVMAAAIPELM--QEDVQIVLLGTGKKFELKLSMEEKYPGKVRVAVVFNAPLAHLIMAG 474

QY 457 ADEMLVPSRFPCCGLIQLHAMHYGTVPVVASTGGGLVDTVIEGTFHMGRLSVDCVKVVEP 514
DB 475 ADVLAVPSRFPCCGLIQLQGRYGTIPCACTGGGLVDTVIEGTFHMGRLSVDCVKVVEP 534

QY 515 ADADALAATVRRASEVFAAGGRYPENVANCISODLSWSKPAQKWEGLLEEVYVYKGG---- 570
DB 535 SDVKKVAATLKRAIKVGTGPAYEEMVRNOMDLSWKGPKNWENVL--LGLGVAGSAPG 592

QY 571 -----VATAKKEIKVP 582
DB 593 IEGDEIAPLAKENVAAP 609

```

[illegible]

FT CHAIN 77 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
 FT BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 608 AA; 67000 MW; CCA8FA50A2F69CB0 CRC64;

Query Match
 Best Local Similarity 51.9%; Score 1417; DB 1; Length 608;
 Matches 299; Conservative 62; Mismatches 171; Indels 44; Gaps 11;

QY 32 LRELARGARKSTSSAVTGATC-----ALDIVMVAEAVAPWSKTGGLGDTGGLPIEL 88
 DB 52 LRTSARLAKMEGRKRVQWQAGTIVCKQGMNLMVFVGCBEQWCKTGGGLGVLGPPAL 111

QY 89 VKRGHRVMTIAPRYDQYADAWDTSVVD-IMG---EKVRYFHSIKGVHRVWIDHPWFLA 144
 DB 112 AARGHRVMTIAPRYDQYADAWDTSVVD-IMG---EKVRYFHSIKGVHRVWIDHPWFLA 171

QY 145 KVMGTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLP-----GP-GEDEVFA 196
 DB 172 KVMGTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLP-----GP-GEDEVFA 231

QY 197 NDHSAVPLVLLKDYQPKQGTAKSVLAHINIAFOGRMWEAEAKDTKLPQAAFKLAF 256
 DB 232 NDHSAVPLVLLKDYQPKQGTAKSVLAHINIAFOGRMWEAEAKDTKLPQAAFKLAF 291

QY 257 SDGYAKVYTEATPMEDEKPPLTGTYKINWLGKGIITAAADKLTVSPNVATEIAADAAG 316
 DB 292 IDGYDK-----PVKG---RKINWKGAGIETADRVFVSPNAYKELVSCVSK 334

QY 317 GVLELTVIRAKGIEGIVMGDIETENPKTKFSLAPYDONSIVYAGKAAKALQAEGLP 376
 DB 335 GVLELTVIRAKGIEGIVMGDIETENPKTKFSLAPYDONSIVYAGKAAKALQAEGLP 394

QY 377 VDPTAPLFAFGRLEEQKGVDIILALPKILATPKVQITAILGTGAAYEKLNAIGTKYK 436
 DB 395 VDRNPLPLGIFGRLEEQKGVDIILALPKILATPKVQITAILGTGAAYEKLNAIGTKYK 453

QY 437 GRAKGVKFSAPLAHMLTAGADFMVPSRFPCGLIQLHAMHYGTVPVASTGGGLVDTVK 496
 DB 454 DKARGVAFNVPPLAMITAGADFMVPSRFPCGLIQLHAMHYGTVPVASTGGGLVDTVK 513

QY 497 EGVTFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCISODLSWSKPA 554
 DB 514 EGVTFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCISODLSWSKPA 573

QY 555 QKWESLL-----EYVYKGGVATAKKEIKVP 582
 DB 574 KMWETVLLSLGVAGSEPGVEGE-EIAPLAKENVATP 608

RESULT 9

UGST_SORBI
 ID UGST_SORBI STANDARD; PRT; 608 AA.
 AC Q43134;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Granule-bound glycogen [starch] synthase, chloroplast precursor
 DE (EC 2.4.1.11).
 GN WAXY OR WX.
 OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Sorghum.
 OX NCBI_TaxID=4538;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. 12311; TISSUE=Seed;
 RA Hsing Y.C., Liu C., Yu H., Hsieh J.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
 CC UDP + [(1,4)-alpha-D-glucosyl](N+1).
 CC -!- PATHWAY: Starch biosynthesis.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U23945; AAC49804.1; -;
 DR InterPro; IPR001296; Glycos_transf_1.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 KW Glycogen biosynthesis; transferase; Glycosyltransferase;
 FT Transit peptide; Chloroplast; Starch biosynthesis.
 FT TRANSIT 1 77 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 78 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
 FT BINDING 97 97 UDP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 608 AA; 66074 MW; C31333FA87D2D8A6 CRC64;

Query Match 37.8%; Score 1376; DB 1; Length 608;
 Best Local Similarity 46.7%; Pred. No. 3.1e-74;
 Matches 295; Conservative 83; Mismatches 180; Indels 74; Gaps 14;

QY 1 MAVASTSRPSSARPPIVINAASGV-----KKTANQLLRELARGSKSTSRSAVTGA-- 52
 DB 1 MSTLATSQ-----LVATHAGLVDPASMFRRGGVQGLRAAARASAAAGDALSMRTSACP 54

QY 53 -----TCATCALDIVMVAEAVAPWSKTGGLGDTGGLPIELVKR 91
 DB 55 AFRQQAARRGGRRGRRPPLVVCATAGMNVFVGAEMAPWSKTGGLGDLGLPAPMAAN 114

QY 92 GHRVMTIAPRYDQYADAWDTSVVVDI-MG---EKVRYFHSIKGVHRVWIDHPWFLAKV 147
 DB 115 GHRVMTIAPRYDQYADAWDTSVVSEIKMGDGYETVRFHCYKGVDRVIDHPLFLERW 174

QY 148 GKTGSKLYGPRGADYLDNHRKFALECKAAIEAARVLP-----PF--GP-GEDEVFA 199
 DB 175 GKTEKIYGPDACTDYKDNQLRFLSLCQAALAPRILSLNNPFFSGPYGEDVVFVCNDW 234

QY 200 HSALVPLVLLKDYQPKQGTAKSVLAHINIAFOGRMWEAEAKDTKLPQAAFKLAFSDG 259
 DB 235 HTGPLSCYLKSNFQSNIGYKDAKTAFCIHNIISQGRFAFSDFPELNLPERFKSSPFD 294

QY 260 YAKVYTEATPMEDEKPPLTGTYKINWLGKGIITAAADKLTVSPNVATEIAADAAGVE 319
 DB 295 YEK-----PVGE---RKINWKGAGIETADRVFVSPNAYKELVSCVSK 337

QY 320 LDTVIRAKGIEGIVMGDIETENPKTKFSLAPYDONSIVYAGKAAKALQAEGLPVPD 379
 DB 338 LDNIMRLTGITGIVMGDMVSEWDPKDYIAVQVSTAVEAKALKEALQAEVGLPVD 397

QY 380 TAPLEAFGRLEEQKGVDIILALPKILATPKVQITAILGTGAAYEKLNAIGTKYKGA 439
 DB 398 KIPLVAFGRLEEQKGVDMVAAAIP-LLMEEDIQIVLLGTGKKKFERMLMSAEKYPDKV 456

QY 440 KGVKFSAPLAHMLTAGADFMVPSRFPCGLIQLHAMHYGTVPVASTGGGLVDTVKGV 499
 DB 457 RAVVKNALAAHIMAGADLLAVTSFEPGLIQLQGMRYGTFPCASTGGGLVDTVIEK 516

QY 500 TGFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCISODLSWSKPAQK 557
 DB 517 TGFHMGALNPD--KLDEADADALAAATVRRASEVFAAGRYPEMVANCISODLSWSKPAQK 576

QY 558 EGLLEE--VYKGG-----VATAKKEIKVP 582
 DB 577 ENVLVSLGVAGSEPGVEGE-EIAPLAKENVATP 608

RESULT 10

UGST_WHEAT
 ID UGST_WHEAT STANDARD; PRT; 615 AA.

QY	240	AFKDTKLPQAFPKLAFSDGYAKVYVTEATPWEDEKEPPLTKGTYKKINWLKGGIIADKL	299
Db	281	DFAQLNLPDRFKSDFIDYDK-----PVGE---RKINWKGAGLQADKV	323
QY	300	VTPSPNATETIADAAGVGVELDVIRAKGTGEGIVGMGDIENWPKTDKELSPYDQNSVY	359
Db	324	LTVSPYAEELISCEARGCELDNIMRLTGIVGMGDMVSEWDPIKOKFLIVNDVTAL	383
QY	360	AGRAAAKEALQAEGLPVDPTAPLFAFIFGRLEEKGVDIILAAALPKILATPKVQJALIGT	419
Db	384	EGRALKEALQAEVGLPDRKVPPLVAFIFGRLEEKGVDVMAIAPEIVKEEDVQIVLIGT	443
QY	420	GKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADPMLVPSRPEPCGLIOLHAMHY	479
Db	444	GKKKFERLLKSVSEKFTKVRVVRFPAPLAHQMMAGADVLAVTSREPECGLIOLQGMRY	503
QY	480	GTPVFWASTGGLVDIVKEGVTGFHMGALNPD--KLDEADADALAAATVRRASEVEFAGGRYP	537
Db	504	GTPCACASTGGLVDIVKEGVTGFHMGRLSVDNCNVVEPADVKVVTTLKRAVKVVGTPAYH	563
QY	538	EMVANCISQDLSSKPAQKWEGLLEEV	564
Db	564	EMVKNCMIODLSWKGPAKNWEDVLEL	590
RESULT 11			
UGST_MAIZE			
ID	UGST_MAIZE	STANDARD;	PRT; 605 AA.
AC	P04713;		
DT	13-AUG-1987 (Rel. 05, Created)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Granule-bound glycogen [starch] synthase, chloroplast precursor		
DE	(EC 2.4.1.11).		
GN	WAXY.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;		
OC	Panicoidae; Andropogoneae; Zea.		
OX	NCBI_TaxID=4577;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RA	Kloegen R.B., Gierl A., Schwarz-Sommer Z., Saedler H.;		
RT	"Molecular analysis of the waxy locus of Zea mays.";		
RL	Mol. Gen. Genet. 203:237-244(1986).		
CC	-1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.		
CC	-1- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =		
CC	UDP + [(1,4)-alpha-D-glucosyl](N+1).		
CC	-1- PATHWAY: Starch biosynthesis.		
CC	-1- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.		
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE		
CC	FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)		
CC	or send an email to license@isb-sib.ch .		
CC	-----		
DR	EMBL; X03935; CAA27574.1; -;		
DR	EMBL; M24258; AAA33520.1; -;		
DR	PIR; S07314; S07314.		
DR	MaizeDB; 15806; -;		
DR	InterPro; IPR001296; Glycos_transf_1.		
DR	Pfam; PF00534; Glycos_transf_1; 1.		
KW	Glycogen biosynthesis; Transferase; Glycosyltransferase;		
KW	transit peptide; Chloroplast; Starch biosynthesis.		
FT	TRANSIT 1 72 CHLOROPLAST.		
FT	CHAIN 73 605 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.		
FT	BINDING 91 991 UDP-GLUCOSE (BY SIMILARITY).		


```
SQ SEQUENCE 605 AA; 65966 MW; 137F15207DBFC189 CRC64;
Query Match 37.6%; Score 1369.5; DB 1; Length 605;
Best Local Similarity 46.7%; Pred. No. 7.5e-74;
Matches 294; Conservative 87; Mismatches 177; Indels 71; Gaps 16;

QY 1 MAVATSRPSSARPIVINAASFGVKTKTANQLLRELARG--SARKSTSRSAVTGATGA--- 55
DB 1 MAALATSQ-----LVATRAGLVG-PDASTFRGAAGLGRASAAADTSLMSRTSRAAA 53
QY 56 -----TCA---LDIVMVAEAPWPKTGLGVDGVTGGLPGLVLELKRGRHV 95
DB 54 PRHQOQARRGRFSLVVCASAGNMVVFVGAEMAPWPKTGLGVDGVLGGLPAMANGHRV 113
QY 96 MTIAPRDQADAWDTSVVVDI-MG---EKVRYFHSIKKKVHRVWIDHPWFLAKVWGKTG 151
DB 114 MVVSPRDQDKAWDTSVVSEIKMGDGYETVRFFCYKRGVDRVFDVDPFLERVWGKTE 173
QY 152 SKLYGPRSGADYLDNHRKFALECKAAFEAAVRL-----PF--GP-GEDCVFVANDHSA 203
DB 174 EKIVGPVAGTYDRNQRLRFLSCQALEAPRLILSLNNPYPFSGYGEDVFEVCDNHTGP 233
QY 204 VPVLKDEYOPKQFTKAKSVLAIHNAFQGRMWEAAFKDTKLPQAAFDKLAIFSDGYAKV 263
DB 234 LSCYLKSNYQSHGIYRDAKTAFCIHNIISYQGRFAFSYPELNLPERKSSDFDIDYEK- 292
QY 264 YTEATPDEEKPLTKTKYKINWLKGGIIADKLVTVSPNYATEITAADAAGVELDTV 323
DB 293 -----PVGE---RKINMKAGILEADRVLTVPYPAEELISGIARGCELDNI 336
QY 324 IRAKGIEGVNGMDIEWNPKTDFLSAPYDONSVYAGKAAKEALQALQELGLPDPAPL 383
DB 337 MRLTGITGVNGMDVSEWDSRDKYIAVKDVSTAVAKALKQALQAEVGLPVDRIPL 396
QY 384 FAFGRLEQKGVDIILAALPKIL-ATPKVQIAILGTGKAAEYKLVNAIGTKYKGRAGV 442
DB 397 VAFGRLEQKGVDMVMAAIPQLMEMVEDYQVILGTGKKKFERMMSAEKFFGKVRV 456
QY 443 VKFSAPLAHMLTAGADPMVPSRPEPCGLIQLHAMHYTPVPVASTGGLVTVKGVYTF 502
DB 457 VKNAAALAHMAGADVLAVTSRPEPCGLIQLGMRGTGTCACASTGGLVTVIEGKTGF 516
QY 503 HMGALNPD--KLDEADADALAATVRRASEVPAGSRYPBMVANCISQDLSNKSQAQKEGL 560
DB 517 HMGELSDVCNVPEADPVKKVATLQRAIKVGVTPAYEEMVRNCMIQDLSMKGPAAKNW 576
QY 561 LEE--VYVGKGV-----ATAKKEEIKVP 582
DB 577 LLSLGVAGGEPGVGEEETIAPLAKENVAAP 605

RESULT 12
ID UGS2_SOLTU STANDARD; PRT; 641 AA.
AC P93568;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Soluble glycogen (starch) synthase, chloroplast precursor
DE (EC 2.4.1.1) (SS I).
OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Desiree; TISSUE=Leaf;
RX MEDLINE=97164391; PubMed=9011082;
RA Abel G.J.W., Springer F., Willmitzer L., Kossmann J.;
RT "Cloning and functional analysis of a cDNA encoding a novel 139 kDa
RL starch synthase from potato (Solanum tuberosum L.).";
plant J. 10:981-991(1996).
```

```
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -!- PATHWAY: Starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y10416; CAA71442.1; -
CC InterPro: IPR001296; Glycos_transf_1.
CC Pfam: PF00534; Glycos_transf_1; 1.
CC KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
CC FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC CHAIN 1 ? 641 SOLUBLE GLYCOGEN (STARCH) SYNTHASE.
CC FT BINDING 145 145 UDP-GLUCOSE (BY SIMILARITY).
CC SQ SEQUENCE 641 AA; 70608 MW; 30F8B0546CEFB74C CRC64;
Query Match 24.8%; Score 903; DB 1; Length 641;
Best Local Similarity 38.1%; Pred. No. 3.1e-46;
Matches 219; Conservative 91; Mismatches 183; Indels 82; Gaps 20;
QY 30 QLLRELARGSARKSTSRSAVTG-----ATGATCALDIVMVAEAPWPKTGG 76
DB 89 QLIPHSVAGDATMVESHEDIVANDRDLSEDTEEMETPIKLTFTIIFVTAAPYKSG 148
QY 77 LGDVTGGPIELVKGHRVMTIAPRY-----DOYADAWDTSV--VVDIMG--EKVYTF 125
DB 149 LGDVGSLPMAALARGHRVMSVSPRYLNGGSPDEKYANAVDLDRATVHCFGDAQEAVF 208
QY 126 HSIKGRVHRVWIDHPWFLAKVWGKTGSKLYGPRSCADYLDNHRKFALECKAAIEAARVLP 185
DB 209 HEYRAGVDWVFDHSSY-----RPGTP-YGDIYGA-FGDNQFRTLLSHACAPVLP 261
QY 186 FGP--GDCVFNANDHMSALVPVLLKDEYOPKQFTKAKSVLAIHNAFQGRMWEAAFK 242
DB 262 LGGFTYGEKCLFLANDHAAALVPLLLAAKRYPYGVYKDARSIVAHNIAHQGVPEAVTN 321
QY 243 DTKLPQAAFDKLAIFSDGYAKVYVTEATPMEDEKPLTKTKYKINWLKGGIIADKLVTV 302
DB 322 NLGLPQWYGVAV---EWIFPTWARAHALD-----TGET---VNVLKGAIVADRIITV 368
QY 303 SPNYATEIAADAAGVELDTVI--RAKGIEGVNGMDIEWNPKTDFLSAPYDONSVA 360
DB 369 SOGYSWETIT-PEGGYGLHELLSSRSQSVLNGITNGIDVNDWNPSTDEHIAHSHYNDL-S 426
QY 361 GKAAAKEALQALQELGLPDPAPLFAFGRLEQKGVLDIILAALPKIILATPKVQIAILGTG 420
DB 427 GKVOCKTDLQELGLPIRPPCPLIGFTGRLDYQKGVLDIILAALQKQ--DVQVVMLGSG 485
QY 421 KAAEYKLVNAIGTKYKGRKGVVKSAPLAHMLTAGADEMLVPSRFPCEGLIQLHAMHYG 480
DB 486 EKQYEDWHRHTENLFKDKFRWVGFNVPVSHRITAGCDIILMPSRFPCEGLNQIYAMRYG 545
QY 481 TVPVPVASTGGLVDTVK-----EGV---TGFMHGMALNPKDLDEADADALAATAVRRASEV 530
DB 546 TIPIVHSTGGLRDTVKDFNPNYQAGGIEGEGTGWTFSPLTSEKL-----LDTLKLAI----- 595
QY 531 FAGGRYPE-----MVANCISQDLSWSKPAQKWE 558
DB 596 ---GTYTEHKSSWEGLMRGRMDRYSWENAAIQVE 627
RESULT 13
UGS3_PEA
```

QY	294	IADKLVTVSPNVAETEATAADAAGVELDVTIRAKG--IEGIVNGMDIEEWNPKTDKFLSA	351
Dd	474	KTADRIVTVSHGVAYELKT-SEGSGWGLHNTINESDKMFKRGIIVNGVDTKOMNQFDAYLTS	532
QY	352	P---YDNSSVYAGKAARKEALCAQLGLVPDPAPLFAFTIGRLSEOKGVDIIILAAALPKIL	407
Dd	533	DGYTNVLKTLQTGRCKROCKAALRELGLPVEDVPPIISFGLRDHKDVGVDLIAEAIPWM	592
QY	408	ATPKVQIAIILGTGKAAYEKILVNAGTKYKGRAGVVKFSAPLAHMITAGADFMLVPSRPE	467
Dd	593	-SHDVQLVMJLTGRADLEQMKEFEAQHCDCIKRSWVGVSFKMAHRITAGSDIILMPSRFE	651
QY	468	PCGLIOLHAMHYGTVPVAVSTGGVLDTVK-----EGVTGFHMGALNPDKLDADADALA	521
Dd	652	PCGLNQDIAMSYGTVPVVHGVGVGLRDVTQVFENFDESGVGW-----TFDRAEANKLM	703
QY	522	ATVRRASEVPAGGRYPGMVANC-----SODLSWSKPAQKWEGGLEEE	563
Dd	704	AA-----LNCLLIITYKKYSWEGIQERGMSODLSWDNAAOQEYLVA	747
QY	564	VVY 566	
Dd	748	AKY 750	
 RESULT 14 UGS2_ORYSA			
ID	UGS2_ORYSA	STANDARD;	PRT; 626 AA.
AC	Q40739;		
CT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Soluble glycogen [starch] synthase, chloroplast precursor		
DE	(EC 2.4.1.11) (SSS).		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartideae; Oryzeae; Oryza.		
OX	NCBI_TaxID=4530;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.		
RC	STRAIN=cv. Japonica; TISSUE=Seed;		
RX	MEDLINE=94302151; PubMed=7518089;		
RA	Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,		
RA	Kobayashi E., Ohnishi S., Tanaka K.-I., Arai Y.;		
RT	"Identification, cDNA cloning, and gene expression of soluble starch		
RT	synthase in rice (<i>Oryza sativa</i> L.) immature seeds.";		
RL	Plant Physiol. 103:565-573(1993)..		
CC	-!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =		
CC	UDP + {(1,4)-alpha-D-glucosyl}(N+1).		
CC	-!- PATHWAY: Starch biosynthesis.		
CC	-!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.		
CC	-!- TISSUE SPECIFICITY: LEAVES AND IMMAIURE SEEDS.		
CC	-!- MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE		
CC	PURIFIED: RSS1, RSS2 AND RSS3.		
CC	-!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE		
CC	FAMILY.		

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL: D16202; BAA03739.1; "		
DR	InterPro: IPR001296; Glycos_transf_1.		
DR	Pfam: PF00534; Glycos_transf_1; 1.		
KW	Glycogen biosynthesis; Transferase; Glycosyltransferase;		
KW	Transit peptide; Amyloplast; Amyloplast; Starch biosynthesis.		
FT	TRANSIT 1 113		
FT	CHLOROPLAST (POTENTIAL).		

FT	CHAIN	114	626	SOLUBLE GLYCOCEN [STARCH]	SYNTHASE, RSS3.
FT	CHAIN	122	626	SOLUBLE GLYCOCEN [STARCH]	SYNTHASE, RSS1.
FT	BINDING	147	147	UDP-GLUCOSE (BY SIMILARITY).	
SEQ	SEQUENCE	626 AA;	68451 MW;	03E4182507D26658 CRC64;	

Query Match		24.0%;	Score 873.5;	DB 1;	Length 626;
Best Local Similarity		41.5%;	Pred. No. 1.7e-44;		
Matches 197;		Conservative 77;	Mismatches 154;	Indels 47;	Gaps 13;

QY	61	IVMVAEEVAPWSKTTGGLDVTGGLPIELVKRGHRWTIAPRY-----DOYADAWDTSV	111
Db	135	VVFVTGEASPYAKSGGLGDVCGSLPIALALRGHRVVMVMPRYMNGALNKNFANAFYTEKH	194
QY	115	VDIM--GE-KVRFHSHKGGVHRVTDHPWELAKVWGKTGSKLYGPRSGADYLDNHRKF	170
Db	195	IKPCTGGEHEVTFHFHYRSDVWVFVDHPSY-----HRPGNLYGDNFGA-FGDNGFRY	247
QY	171	ALFCKAAIEAARVLPEFG---GEDCVFVANDHWSALYPLLKDRYQKQGFTKAKSVLAI	227
Db	248	TLLCYAAACEAPLILELGGYIYGKCMFVVDHWSALYPLLAAKYRYPGYVRDARSVLVI	307
QY	228	HNIAFGRMGEAEFKDKTLFOAAPDKLAFSDGYAKVYTEATPMEDEKPLTGTGKYKIN	287
Db	308	HNLAHQGEVPASTYPLDGLPPEWYGALEW-----VPEWARRHADKG-----EAVN	354
QY	288	WLKGGIIAADKLVITSPNYATEAADAAGVDELDTVI--RAKGTEGIVNGMDIEWNPKT	345
Db	355	FLKGVVTAORIVTVSGOYSWEVTT-AEGGGNELLSSRSKSVLNGIDINDNNWNPST	413
QY	346	DKFLSAPYDONSVAAGKAAAEALQAEGLGVPDPTAPLFAFGRLEQKGVDDIILAALPK	405
Db	414	DKELPYHYSVDLL-SGKAKCAELQELGPIRPDPVPLIGFGRKDYQKGLDILTKLAIPD	472
QY	406	ILATPKVQTALGTGKAAYEKLVAIGTKYKRAKGVKVESAPLAHMLTACADFMLVPSR	465
Db	473	LM-RDNIQTQFVMLGSGDPFGEGWMKSTESGYRDKFRGCVGFSVPVSHRITACDILLMPSR	531
QY	466	FEPCCGLIQLHAMHYCHVPVVASGTGGLVDTVK-----EGVTGFHMGALNPDK	511
Db	532	FEPCCNLQIYAMGYGTPVYVHGHTGGRLDVFENFENPAEKGEQGTGWAFSLPTEK	586

```

RESULT 15
UGS2_WHEAT
ID UGS2_WHEAT STANDARD; PRT; 610 AA.
AC Q43654;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Soluble glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.1) (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Triticum.
OC NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. T.A. Florida; TISSUE=Endosperm;
RA Block M., Loerz H., Luetticke S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -1- PATHWAY: Starch biosynthesis.
CC -1- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U48227; AB02197.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
FT NON_TER 1 1
FT TRANSIT <1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN ? 610 SOLUBLE GLYCAGEN [STARCH] SYNTHASE.
FT BINDING 7 7 UDP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 610 AA; 67143 MW; 46508A3BF7B87193 CRC64;

```

Query Match	23.9%	Score 869;	DB 1;	Length 610;
Best Local Similarity	35.0%;	Pred. No. 3e-44;		
Matches	234;	Conservative 94;	Mismatches 222;	Indels 118; Gaps 22;
QY	67	EVAPWSKTGGLGVDVTGGLPIELNKRGRHVMTIAPRY-----DOYADANTSVVVDIM--	118	
Db	1	EAAPYAKSGGLGDCVGLSLPATAAARGHVMVMPRYLNGSSDKNYAKALYAKHILPCF	60	
QY	119	--GEKYRPHSHIKKGVHRVWDHPWFLAKVMGKTGSKLYPRSGADYLDNHKRFALFCKA	176	
Db	61	GGSHVTFHFHEYRDNDVDVFDHP-----SYHRPGS-LYDNTGA-FGDNQFRYTLLEYA	113	
QY	177	AIFAARVLPGP---GEDCVFVANDHSAIVPVLVKDEYQKQFTKAKSVLAHNTAFQ	233	
Db	114	ACEAPLLELGGYTYGONCFVNDHMASLVPVLLAAKYPGVYRDSRSTLVHNLAHQ	173	
QY	234	GRWEEAFKDKLPOAADFKLAFSDGYAKYVYTEATPMEDEKPPLTGKTYKKLNWLKGGI	293	
Db	174	GVPEASTYDGLGPPPEWYGALW-----VFPWARHRLDKG-----EAVNFKGAV	220	
QY	294	IAADKLVTGSPNATETAADAAGVGLDVTI--RAKIGIEVNGMDTEENKPTDKFSLA	351	
Db	221	VTADRIVTVSGQYSWEVTT-AEGOGLELLSSRKSVLNGIVNGIDINDNPTDKCLPH	279	
QY	352	PYQNSVYAGKAAKAELQELGVPDPTAPLFAFTGRLEEQKGVDIILAALPKILATPK	411	
Db	280	HYSVDDIL-SGKACKAEQLQELGLPVEDVPLGFTIGRLDYQKGLDLKWAPELM-RED	337	
QY	412	VQIAILGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVLVPSRPPCGL	471	
Db	338	VQFVMLGSGDPFIEGWMRSTESSTKDKFRGWGFSVPVSHRITAGCDILLMPRSFPCGL	397	
QY	472	IQLHAMHYGTPPVVASTGGLVDVTVK-----EGVTGHMGALNPDKLDEADADALAA	532	
Db	398	NQLYAMQYGVPPVVGHTGGRLDVTETFPNPFAGKEBGTGWAFSLPVDKMLWA-----	450	
QY	523	TVPRASEVFAGGRYPFWANCISODLSWSKPAQKWEGLLEVVYVYKGGVAIAKKEIKVP	582	
Db	451	-LEIAMSTFRE-----HKPS--WEGLM-----KRGMTKDH	477	
QY	583	VAEKIPDGLPAVSYAFTNLKPVSVASVEGGAAPKVGTTAPAMGAWRATTPSGPSAAAT	642	
Db	478	TWDHAFSSSTRSSGGSWNTPTSC-----RRGLGRSKCSPSALKTSSSS	522	
QY	643	---PKVITYKPALPAKPAKPTAGLK-LAGEASTTSTSENGAASNGNGCASAKSTSAAPK	698	
Db	523	FRGPEGYPCTLRCPAIVEVOCACLLWFAG-----SRTYDGA-----AAAVTASGGRQ	570	
QY	699	LYSAATR K 706		
Db	571	LQFWGIRK 578		

Search completed: June 4, 2003, 14:53:07
Job time : 15.5065 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:35:26 ; Search time 35.1742 Seconds
(without alignments)
4147.394 Million cell updates/sec

Title: US-09-980-771A-3

Perfect score: 3641

Sequence: 1 MAVASTSRPSPARPVINA.....SASKTSRAKPLVSRATRKSA 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3637	99.9	708	10	O64925	O64925 chlamydomon
2	1476.5	40.6	613	10	Q93YB1	Q93YB1 pisum sativ
3	1454.5	39.9	604	10	Q9FR03	Q9FR03 perilla fru
4	1441	39.6	607	10	Q9ZSQ5	Q9ZSQ5 astragalus
5	1434	39.4	606	10	Q9XIS6	Q9XIS6 phaseolus v
6	1428	39.2	608	10	Q93YD9	Q93YD9 ipomoea bat
7	1426	39.2	609	10	Q94LY7	Q94LY7 oryza sativ
8	1419.5	39.0	603	10	Q8SA49	Q8SA49 hordeum vul
9	1418	38.9	609	10	Q8S9C4	Q8S9C4 oryza sativ
10	1409	38.7	605	10	Q9SXX3	Q9SXX3 triticum ae
11	1408	38.7	605	10	Q9FUT6	Q9FUT6 triticum ae
12	1405.5	38.6	606	10	Q43012	Q43012 oryza sativ
13	1405.5	38.6	610	10	Q9MAQ0	Q9MAQ0 arabidopsis
14	1405	38.6	608	10	Q8VYU1	Q8VYU1 oryza sativ
15	1395.5	38.3	604	10	Q9S7N5	Q9S7N5 triticum ae
16	1394	38.3	605	10	Q9SLS8	Q9SLS8 triticum tu

17	1394	38.3	605	10	Q9SQ51	Q9sq51 aegilops sp
18	1388.5	38.1	599	10	Q9SQ58	Q9sq58 triticum ae
19	1387.5	38.1	604	10	Q9SLS6	Q9sls6 triticum tu
20	1384.5	38.0	605	10	Q9SQ52	Q9sq52 triticum mo
21	1383	38.0	604	10	Q9SLS7	Q9sls7 triticum tu
22	1375	37.8	534	10	Q8W2G8	Q8w2g8 triticum ae
23	1374	37.7	574	10	Q9SYU0	Q9syu0 triticum ae
24	1372.5	37.7	604	10	Q9SXX4	Q9sxx4 triticum ae
25	1370.5	37.6	604	10	Q9SLS9	Q9sls9 triticum tu
26	1268.5	34.8	565	10	Q9XEN9	Q9xen9 triticum ae
27	900.5	24.7	647	10	Q9LEB9	Q9leb9 triticum ae
28	900.5	24.7	647	10	Q9SQH0	Q9sqh0 aegilops ta
29	900.5	24.7	647	10	Q9SQG9	Q9sqg9 triticum ae
30	899.5	24.7	647	10	Q9LEC0	Q9lec0 triticum ae
31	898	24.7	792	10	Q9MAC8	Q9mac8 arabidopsis
32	895.5	24.6	313	10	Q9LKD3	Q9lkd3 vaquelinia
33	894.5	24.6	643	10	Q9M5A3	Q9m5a3 hordeum vul
34	893.5	24.5	313	10	Q9LKE0	Q9lke0 prunus virg
35	893.5	24.5	313	10	Q9FYU8	Q9fyu8 aruncus dio
36	886.5	24.3	313	10	Q9LKE8	Q9lke8 kageneckia
37	886	24.3	576	10	O64926	O64926 chlamydomon
38	885.5	24.3	313	10	Q9LKO1	Q9lkd1 exochorda r
39	882	24.2	630	10	O65365	O65365 ipomoea bat
40	879.5	24.2	313	10	Q9LKE5	Q9lke5 amelanchier
41	879.5	24.2	313	10	Q9LKD6	Q9lkd6 vaquelinia
42	878.5	24.1	313	10	Q9LKE7	Q9lke7 kageneckia
43	878.5	24.1	313	10	Q9LKE1	Q9lke1 prinsepia s
44	877.5	24.1	313	10	Q9LKE6	Q9lke6 malus sarge
45	877.5	24.1	313	10	Q9LKD8	Q9lkd8 rosa multif

ALIGNMENTS

RESULT 1

O64925
ID O64925 PRELIMINARY; PRT; 708 AA.
AC O64925.
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Granule-bound starch synthase I precursor (EC 2.4.1.21).
GN STA2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137C;
RA D'Hulst C., Wattlebled F., Ral J.-P., Abel G.-J., Kossmann J.,
RA Ball S.G.;
RT "Cloning of a cDNA encoding for the GBSSI in the green alga
RT Chlamydomonas reinhardtii."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wattlebled F., Ball S.G., D'Hulst C.;
RT "Granule-bound starch synthase I: A major enzyme involved in the
RT biogenesis of B-crystallites in starch granules."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF06420; AAC17969.3;
DR EMBL; AF433156; AAL28128.1;
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR002114; HPr_Serp_site.
DR Pfam; PF00534; Glycos_transf_1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
KW Glycosyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 57 POTENTIAL.
SQ CHAIN 58 708 GRANULE-BOUND STARCH SYNTHASE I.
SQ SEQUENCE 708 AA; 74623 MW; 7D2A5A07D8606469 CRC64;

Query Match

99.9%; Score 3637; DB 10; Length 708;

```
Best Local Similarity 99.9%; Pred. No. 3.6e-205;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAVASTSPSPARPVINAASFGVKKYKKTANQILLRELARGSARKSTSRSAVTGATGATCALD 60
DB 1 MAVASTSPSPARPVINAASFGVKKYKKTANQILLRELARGSARKSTSRSAVTGATGATCALD 60
QY 61 IMVAAEAPVAPSKTGGGLDVTGGLPIELVKKGRHVRMTIAPRYDQYADAWDTSVVVDIMGE 120
DB 61 IMVAAEAPVAPSKTGGGLDVTGGLPIELVKKGRHVRMTIAPRYDQYADAWDTSVVVDIMGE 120
QY 121 KVRVFSHKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFCKAAIEA 180
DB 121 KVRVFSHKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFCKAAIEA 180
QY 181 ARVLPGEGEDCVFVANDWHSALVPLVLLKDEYQPKGQFTKAKSVLAIHNIAPQGRMWEEA 240
DB 181 ARVLPGEGEDCVFVANDWHSALVPLVLLKDEYQPKGQFTKAKSVLAIHNIAPQGRMWEEA 240
QY 241 FKDTKLPQAAFDKLAFSAGYAKVYTEATPMEDEKPLTGTYYKINWLKGGIIAADKLIV 300
DB 241 FKDTKLPQAAFDKLAFSAGYAKVYTEATPMEDEKPLTGTYYKINWLKGGIIAADKLIV 300
QY 301 TVSPNYATEIAADAAAGGVVELDTVIRAKGIEGVNGMDIEWNPKTKDKLSAPYDONSVA 360
DB 301 TVSPNYATEIAADAAAGGVVELDTVIRAKGIEGVNGMDIEWNPKTKDKLSAPYDONSVA 360
QY 361 GKAAAKEALQELGLPDPPTAPLPAFAGRLBQKGVDDIILALPKILATPKVQIAILGTG 420
DB 361 GKAAAKEALQELGLPDPPTAPLPAFAGRLBQKGVDDIILALPKILATPKVQIAILGTG 420
QY 421 KAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYG 480
DB 421 KAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYG 480
QY 481 TVPVVASTGGLVDTVKEGVTGFHMGALNPDKLDEADALAAATVRRASEVFAGGYPYPMV 540
DB 481 TVPVVASTGGLVDTVKEGVTGFHMGALNPDKLDEADALAAATVRRASEVFAGGYPYPMV 540
QY 541 ANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEIKVPVAKIIPGDLPAVSAPNT 600
DB 541 ANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEIKVPVAKIIPGDLPAVSAPNT 600
QY 601 LKPVASVEGNGAAAPKVTGTAPAMGAWRATTPSGPSAAATPKVTTYPALPATAPKPT 660
DB 601 LKPVASVEGNGAAAPKVTGTAPAMGAWRATTPSGPSAAATPKVTTYPALPATAPKPT 660
QY 661 AGLKLAGASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708
DB 661 AGLKLAGASTTSTSENGAASNGNGASAKTSAAKPLVSAATKSA 708

RESULT 2
Q93YB1 PRELIMINARY; PRT; 613 AA.
AC Q93YB1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Granule-bound starch synthase (EC 2.4.1.11).
GN GBSSI.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Edwards A., Vincken J.P., Visser R., Zeeman S., Smith A.M., Martin C.;
RT "Discrete Forms of amylose are synthesised by isoforms of GBSSI in
pea.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ345045; CAC69955.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR001917; NHtransf_2.
DR Pfam; PF00534; Glycos_transf_1.1
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
KW Glycosyltransferase; Transferase; Transf peptide.
FT TRANSIT 1 85
FT CHAIN 86 613 GRANULE BOUND STARCH SYNTHASE.
SQ SEQUENCE 613 AA; 67626 MW; 5A653461128F97EB CRC64;

Query Match 40.6%; Score 1478.5; DB 10; Length 613;
Best Local Similarity 52.0%; Pred. No. 1.2e-78;
Matches 296; Conservative 75; Mismatches 151; Indels 47; Gaps 11;

QY 39 SARKSTSRSAVTGATGATCALDIVMVAEAPVAPSKTGGGLDVTGGLPIELVKKGRHVRMTI 98
DB 39 SARKSTSRSAVTGATGATCALDIVMVAEAPVAPSKTGGGLDVTGGLPIELVKKGRHVRMTI 98
QY 67 SKKGVQRKNVRPKGLIVCGMNLIIIVGTEVAPVSKTGGGLDVLGGLPPALSANGHRVMIV 126
DB 67 SKKGVQRKNVRPKGLIVCGMNLIIIVGTEVAPVSKTGGGLDVLGGLPPALSANGHRVMIV 126
QY 99 APYDQYADAWDTSVVVDI-----MGEKVRVYFHSIKKGVHRVWIDHPWFLAKVWGKTSK 154
DB 99 APYDQYADAWDTSVVVDI-----MGEKVRVYFHSIKKGVHRVWIDHPWFLAKVWGKTSK 154
QY 127 TPYDQYKDAWDTNVTIEVKGDRTEKVRFFHCFKRGVDRVDFVHPFIELEKVGKGTGTL 186
DB 127 TPYDQYKDAWDTNVTIEVKGDRTEKVRFFHCFKRGVDRVDFVHPFIELEKVGKGTGTL 186
QY 155 YGPRSGADYLDNHRKRFALFCKAAIEAARVLPE-----GP-GEDEVFVANDWHSALV 206
DB 155 YGPRSGADYLDNHRKRFALFCKAAIEAARVLPE-----GP-GEDEVFVANDWHSALV 206
QY 187 YGPAAGDDYQDQLRPSIFCQAALAAARVNLKSNKYSFGPIGEDVIFVANDWHTALISC 246
DB 187 YGPAAGDDYQDQLRPSIFCQAALAAARVNLKSNKYSFGPIGEDVIFVANDWHTALISC 246
QY 207 LKDEYQPKGQFTKAKSVLAIHNIAPQGRMWEEAFKDK---TKLPQAAFDKLAFSAGYAKV 263
DB 207 LKDEYQPKGQFTKAKSVLAIHNIAPQGRMWEEAFKDK---TKLPQAAFDKLAFSAGYAKV 263
QY 247 YNKMVSQSIGIFRNKAVVFCIHNIAYQGRF---AFTDYSLLNLPQFKSSDFLDGHVK- 302
DB 247 YNKMVSQSIGIFRNKAVVFCIHNIAYQGRF---AFTDYSLLNLPQFKSSDFLDGHVK- 302
QY 264 YTEATPMEDEKPLTGTYYKINWLKGGIIAADKLIVTSPNYATEIAADAAAGGVVELDTV 323
DB 264 YTEATPMEDEKPLTGTYYKINWLKGGIIAADKLIVTSPNYATEIAADAAAGGVVELDTV 323
QY 303 -----PIVG---RKINWMKAGIIISHRVLTIVSYQAQLVSPGDKGVDELMI 346
DB 303 -----PIVG---RKINWMKAGIIISHRVLTIVSYQAQLVSPGDKGVDELMI 346
QY 324 IRAKGTIEGVNGMDIEWNPKTKDKLSAPYDONSVAAGKAAKEALQELGLPDPPTAPL 383
DB 324 IRAKGTIEGVNGMDIEWNPKTKDKLSAPYDONSVAAGKAAKEALQELGLPDPPTAPL 383
QY 347 LRRVGTGVINGMDVQENPNSTDKYISIKYDASTVLEGLKALLKEELQAEVGLPDKNVPL 406
DB 347 LRRVGTGVINGMDVQENPNSTDKYISIKYDASTVLEGLKALLKEELQAEVGLPDKNVPL 406
QY 384 FAFIGRLBQKGVDDIILALPKILATPKVQIAILGTGKAAEKLVAIGTKYKGRAGVY 443
DB 384 FAFIGRLBQKGVDDIILALPKILATPKVQIAILGTGKAAEKLVAIGTKYKGRAGVY 443
QY 407 IAFIGRLBQKGVDDIILALPKILATPKVQIAILGTGKAAEKLVAIGTKYKGRAGVY 465
DB 407 IAFIGRLBQKGVDDIILALPKILATPKVQIAILGTGKAAEKLVAIGTKYKGRAGVY 465
QY 444 KESAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYGTVPVVASTGGLVDTVKEGVTGFH 503
DB 444 KESAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYGTVPVVASTGGLVDTVKEGVTGFH 503
QY 466 KENVPLAHMIAAGADFILPSRFEPCGLIQLAMRYGTVPVIVASTGGLVDTVKEGVTGFH 525
DB 466 KENVPLAHMIAAGADFILPSRFEPCGLIQLAMRYGTVPVIVASTGGLVDTVKEGVTGFH 525
QY 504 MGALNP--DKLDEADALAAATVRRASEVFAGGYPYPMVANCISQDLSWSKPAQKWEGLL 561
DB 504 MGALNP--DKLDEADALAAATVRRASEVFAGGYPYPMVANCISQDLSWSKPAQKWEGLL 561
QY 526 MGSFNVKCDADVPDVPDVAIPKTVTKALGVYGTSAFAEMIKMCMQAQELSNKGPAAKWEVL 585
DB 526 MGSFNVKCDADVPDVPDVAIPKTVTKALGVYGTSAFAEMIKMCMQAQELSNKGPAAKWEVL 585
QY 562 -----EEVYKGGVATAKKEIKVP 582
DB 562 -----EEVYKGGVATAKKEIKVP 582
QY 586 LNLGVPDSEFGIDGO-EIAPQAKENVATP 613
DB 586 LNLGVPDSEFGIDGO-EIAPQAKENVATP 613

RESULT 3
Q9FR03 PRELIMINARY; PRT; 604 AA.
AC Q9FR03;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Granule-bound starch synthase.
GN GBSSI.
OS Perilla frutescens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
OX NCBI_TaxID=48386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. OKDONG; TISSUE=DEVELOPING SEEDS;
RA Hwang S.-K., Hwang Y.-S.;
```



```

OC Ehrhartoideae: Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MILKY QUEEN;
RA Sato H., Suzuki Y., Sakai M., Imbe T.;
RT "Molecular Characterization of wx-1, a Novel Mutant Gene for Low-
RT amylose Content in Endosperm of Rice (Oryza sativa L.).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB066093; BAB86209.1; -.
SQ SEQUENCE 609 AA; 66431 MW; B57CC13E0440D227 CRC64;

Query Match      38.9%; Score 1418; DB 10; Length 609;
Best Local Similarity 48.8%; Pred. No. 4.2e-75;
Matches 301; Conservative 81; Mismatches 173; Indels 62; Gaps 13;

QY 3 VASTSRPS-----ARPIVINAASGVKKTAN---QLLRELARGSARKTSRSA 48
Db 18 IADRSAPSSLLRHGFQGLKPRSPAGGDATSLSVTSARATPKQQRSVQGRSR--FPSV 74

QY 49 VTGATGATCADDIVMVAEAPWSKTGGGLGVDVTGGLPIELVKRHRVMTIAPRYQYADA 108
Db 75 VYVATGA--GMNVVVGEMAPWSKTGGGLGVDLGLPPAMAANGHRVWISPRYQYKDA 132

QY 109 WDTVVVDI----MGEKVYRHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGAYL 164
Db 133 WDTVVVAEIKVADRYERVRFFCHYKHGVDVRFIDHPSFLEKVGWKTGEKIYGPDTGV 192

QY 165 DNHKFALFEKAAIEAARVL-----PFGP--GEDCVFVANDHWSALVPLKDEYOPKG 216
Db 193 DNOMFSLCQAALAPRIILNNNPYFKGTGEDVVFVNDWHTGTLASYLKNNYOPNG 252

QY 217 OFTKAKSVLAHINIAFQGRMWEAFKDKLPQAAFDKLAFLSDGAKVYTEATPMEDEKP 276
Db 253 IYRNKAVFACIHNISYQGFEDYPELNLSEFRSSPFDIDY-----DTPVEG--- 302

QY 277 PLTGKTKYKINNLKGGITAAADLVTVSPNYATEIAADAAGGVLELDTVIRAKGIEGIVN 336
Db 303 -----RKINNMKAGILEADRLAVVSPYTABELISGARGCELDNIMRLTIGTIVNGM 355

QY 337 DIEEWNPKDKLSAPYDONSVYAGKAAKEALQALQELGVPDPTAPLFAFGRLEPKGV 396
Db 356 DVSEWDPKDKYITAKYDATTAEAKALNKEALQALQAGLPVDKIPLIAPFGRLEQKGP 415

QY 397 DIILALPKILATPKVQIAILGTGKAAEYKLVNAIGTKYGRAGKGVKFSAPLAHMLTAG 456
Db 416 DWMAAIPELM-QEDVQIVLLGTGKKFKELKSMEEKYPGKVRVAVVKNAPLAHLIMAG 474

QY 457 ADFMLVPSRFEPCGLLIQLHAMHYGVVAVSTGGGLVDTVKEGVTGFHMGALNPD--KLDE 514
Db 475 ADVLAVPSRFEPCGLLIQLQCMRYGTPCACASTGGGLVDTVIEGTFHMGRLSYDCKVVEP 534

QY 515 ADADALAAVVRASEVAFGRYPPEMVANCISODLSWSPKPAQKWEGLLEEVVYKGG---- 570
Db 535 SDVKVKAATLTKRAIKVVGTPPAVEEMVRNMCNQLDSWKGPAKNWENVL--LGLGVAGSAPG 592

QY 571 -----VATAKKEIKVP 582
Db 593 IEGDEIAPLAKENVAAP 609

RESULT 10
Q9SXK3 PRELIMINARY; PRT; 605 AA.
AC Q9SXK3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Starch synthase (GBSSI) (EC 2.4.1.21).
DN WAXY.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;

```

```

OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99321800; PubMed=10393240;
RA Murai J., Taira T., Ohta D.;
RT "Isolation and characterization of the three Waxy genes encoding the
RT granule-bound starch synthase in hexaploid wheat.";
RL Gene 234:71-79(1999).
DR EMBL; AB019623; BAA77351.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00334; Glycos_transf_1; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 605 AA; 66345 MW; 60A816276178722B CRC64;

Query Match      38.7%; Score 1409; DB 10; Length 605;
Best Local Similarity 49.4%; Pred. No. 1.4e-74;
Matches 295; Conservative 77; Mismatches 181; Indels 44; Gaps 11;

QY 8 RPSARPIVINAASFGVKKTANQLLRELARGSARKTSRSVATGATCATCADDIVMVAE 67
Db 31 RPSPADAPLGMRTTGASAAAPKQQRKAHRGTRR---CLSMVVVRATG-SAGMNLVFGAE 86

QY 68 VAPWSKTGGGLGVDVTGGLPIELVKRHRVMTIAPRYQYADAWDTSVVVDIM---GEKVR 123
Db 87 MAPWSKTGGGLGVDLGLPPAMAANGHRVWISPRYQYKDAWDTSVVSEIKVADEYVR 146

QY 124 YHSHIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRFALFCFAAIEAARV 183
Db 147 YFHCYKGVDRVDFVDFHPCFLEKVGKTKETKIGPDAGTDYEDNQLRSLCQAALAPRI 206

QY 184 L-----PF--GP--GEDCVFVANDHWSALVPLVLDKEYQPKQGTAKSVLAHINIAFQGR 235
Db 207 LDLNPNPYSGPYGVEDVVFVNDWHTGTLACLYLKSNTQSSGIVRTAKVAPCIHNISYQGR 266

QY 236 MWEAFKDKLPQAAFDKLAFLSDGAKVYTEATPMEDEKPPLTGKTKYKINNLKGGIIA 295
Db 267 FSPDDFAQLNLPDRFKSSDFIDGYDK-----PVEG---RKINNMKAGILQ 309

QY 296 ADKLTVTVSPNYATEIAADAAGGVLELDTVIRAKGIEGIVNMDTEENWPKTKDFLSAPYDQ 355
Db 310 ADKLTVTVSPNYATEIAEELISGEARGCELDNIMRLTIGTIVNGMDVSEWDPAKDKFLAANYDV 369

QY 356 NSYTAGKAAKEALQALQELGVPDPTAPLFAFGRLEQKGVDIILALPKILATPKVQIA 415
Db 370 TTALEKALNKEALQAEVGLPVDKVPVAFVIGRLEEQKGPVMAIAIPILKEEDVQIV 429

QY 416 ILGTGKAAEYKLVNAIGTKYGRAGKGVKFSAPLAHMLTAGADFMVPSRFEPCGLLIQLH 475
Db 430 LLGTGKKFELRLKSVEEKFPKVRVAVVRFNAPLAHOMMAGADVLAVTSRFEPCGLLIQLQ 489

QY 476 AMHYGTVPVAVSTGGGLVDTVKEGVTGFHMGALNPD--KLDEADADALAAVVRASEVAFG 533
Db 490 GMYGTGPPACASTGGGLVDTVIEGTFHMGRLSDCNVVEPADVKKVVTILKRAVKVGT 549

QY 534 GRYPPEMVANCISODLSWSPKPAQKWEGLLEEV-----VYGGGVATAKKEIKVP 582
Db 550 PAYHEMVKNCMIQDLSWKGPAKNWEDVILELVEGSEPPGVIQE-EIAPLAMENVAAP 605

RESULT 11
Q9FUU6 PRELIMINARY; PRT; 605 AA.
AC Q9FUU6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Granule bound starch synthase I (EC 2.4.1.21).
DN GBSSI.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Triticum.
 OC NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CHEYENNE; TISSUE=ENDOSPERM;
 RA McCue K.F., Hurkman W.J., Tanaka C.K., Anderson O.D.;
 RT "Starch Branching Enzymes Sbel and Sbe2 from Wheat (*Triticum aestivum*
 RT cv. Cheyenne): Molecular Characterization, Developmental Expression,
 RT and Homolog Assignment by Differential PCR.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF286320; AAG27624.1; -;
 DR InterPro; IPR001296; Glycos_transf_1.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
 DR GlycoSite; Transferase; Transferase.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 605 AA; 66326 MW; 21120D9D6F68B891 CRC64;

Query Match 38.7%; Score 1408; DB 10; Length 605;
 Best Local Similarity 49.4%; Pred. No. 1.6e-74;
 Matches 295; Conservative 77; Mismatches 181; Indels 44; Gaps 11;

QY 8 RPSSARPITVNAAGSGVKKTANQLLRRLARGSARKSTSRGAVTGATGATCALDIDVMAAE 67
 DB 31 RPRSPADAPLGMRTTGASAAKPKQSRKAHRTTR---CLSNVVRATG-SAGMNLVFFVGA 86
 QY 68 VAPWSKTGGGLDVTGGPIELVKRGHKVMTIAPRYDQYADAWDTISVVVDIM---GEKVR 123
 DB 87 MAPWSKTGGGLDVLGGLPFAKAANGHVMVISPRYDQYKDAWDTISVVSEIKVADEYER 146
 QY 124 YFHSIKGVRHVRVLDHPFLAKYWGKTKLYGPRSGADYLDNKHKRFALFCKAAFEARV 183
 DB 147 YFHCYKSGVDVFDVHPCFLKYGKTKELYGPDAGTDYEDNQLRSLCQAALAPRI 206
 QY 184 L-----PF--GP-GECDVVFVANDHWSALVPLVLKDEYQPKQGTAKAKSLAIHNIAPQGR 235
 DB 207 LDLNNNFYSPGPGYEDVVFVNCNDWHTGLLACYLKSNYQSSGIYRTAKVAFCHINSYQGR 266
 QY 236 MWEEAFKDTKLQAAAFDKLAFSDGYAKVYTTAATPMEDEKPPLTGKTYKKNLKGKITA 295
 DB 267 FSDDFFAQNLNLPDRKSSPFDIDYDK-----PVVEG---RKNNMKAGILQ 309
 QY 296 ADKLVTVPSPNATETIAADAAGGVELDTVIRAKGLGIVGVNGMDTEENPKTDKFLSAPYDQ 355
 DB 310 ADKVLTVSPYYAEELISGEARGCELDNIMELTGITGVNGMDVSEMDPAKDFLAANYDV 369
 QY 356 NSVTAGKAAAKEALQELGLPVDPTPLFAPFGRLEEQKGVDTILAALPKILATPKVQIA 415
 DB 370 TTALEGAUKNEALQAEVLGVDPVRKVPFLVAFIGRLEEQKGPDMVIAAIPELLKEEDVQIV 429
 QY 416 ILGTGAAAYEKLVAINTGTYKGRAGKVKFSAFLAHMLTAGADFMLVSPFEPGGLIQH 475
 DB 430 LLGTGKKFERLLKSVEEKFPSPKRAVVRFNAPLAHOMMAGADVLVATSRFEPGGLIQ 489
 QY 476 AMHYGVTPVASTGGVLVDVTKGVTFGHMGALNPD--KLDEADADALAAATVRRASVFAG 533
 DB 490 GMRYGTPPCASTGGVLVDITMEGKTGFHMGHLSVDCNVVEPADVKKKVVTTLKRAVKVGT 549
 QY 534 GRYPENVMANICISDLSWSKPAQKWEGLLEEV-----VYCKGGVATAKKKEIKVP 582
 DB 550 PAYHEMVKNCMTQDLSWGPANWEDVLELGVGESEPGVIGE-ETIAPLAMENVAAP 605

RESULT 12	
Q43012	
ID	Q43012
AC	Q43012;
DT	01-NOV-1996 (TREMBlrel. 01, Created)
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE	Starch granule-bound starch synthase (EC 2.4.1.11).
GN	WAXY.

OS *Oryza sativa* (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzaceae; *Oryza*.
OX NCBI_TaxID=4530;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=SSP. L.F. SPONTANEA;
RC Wang Z.Y., Zheng F.Q., Gao J.P., Wang X.Q., Wu M., Zhang J.L.,
RA Hong M.A.;
RT "Identification of two transposon-like elements in rice Wx gene.";
RL Sci. China B 37:437-447(1994).
RR EMBL; X64108; CAA45472.1; --
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR000794; Ketoacyl_synth.
DR Pfam; PF00534; Glycos_transf1_1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 606 AA; 66323 MW; E1157469A553CE CRC64;

Query Match	38.6%;	Score	1405.5;	DB	10;	Length	606;
Best Local Similarity	48.4%;	Pred.	No. 2.3e-74;				
Matches	298;	Conservative	83;	Mismatches	172;	Indels	63;
Gaps	14						
QY	3	VASTSRPSS-----	-----	APPIVINAASFGVYKKTAN	---	QLLREIARGSAKSTSRSA	48
Db	18	IADRGA	PSLLRHGFQGLKPRSPAGDASSLSVTTSARATPKQORSQGRSRR---	FPVS	74		
QY	49	VTCATGATCALDIVMVAAEVAPWSKTGGLGVDVGTGGLPLELWFKGRHRYMTAPRVDYADA	108				
Db	75	VYATGA	---GMNVFVGEAMAPWSKTGGLGDUVLGGLPPMAANGHRRVWVLSPIRYDQTKDA	132			
QY	109	WDTSVVVDI----	MGEKVRYFHSIKKGVHRWIDHPWFLAKVMGKTGSKLYGPRSGADYL	164			
Db	133	WDTSVAAEIKVADRYERVRFHCYKRGVDRVFI	DHPSFLEKVMGKTGKIYGDVTGVYK	192			
QY	165	DNHKRFALFCFAAIEAARVL-----	PTGP---	GEDCVFVANDHWSALVPVLLKDEYTPQKG	216		
Db	193	DNQMRFSLQC---	EAPRILMLNNPNYFKGYGEDVYFVGVNDWHTGFLPSYLNKNTYQNG	249			
QY	217	QFTKAKSVLAHNTAFQGRMWEAEFKDTKLPQAAFDKLA	FSGQYAKVYTEATPMEDEKP	276			
Db	250	IYRNAKVAFCIHNIISYQGRFAEDY	PELNLSERFKSSFDIDGY-----	DTPEVG----	299		
QY	277	PLTGKTYKINWLKGGIITADKLVTSPNVAIEITAADAAGVVELDTVIRAKLEGIVNGM	336				
Db	300	-----	RKIINWKAGILESORVITVSPYAAEELISGIARGCELDNIMRLTGITGVNGM	352			
QY	337	DTEENPNKTDKFLSAPYDONSIVYAGKAAAEKALQAEGLGDPVDTAPLEAFI	GLREOKGY	396			
Db	353	DYSEWDPSKXIIITTKYDATTA	TEAKALNKEALQAEAGLVPDRKRVPLIAF	IGLREOKGP	412		
QY	397	DLIIAALPKIIATPKVQIAILGTGKAAEYKLVNAIGTKYKGRAGKGVYKFSAPLAHMLTAG	456				
Db	413	DVMAAIPELM-QENVQIVLLGTGKKKFEKLLKSMEEKYPGKYRAVVKFNAPLAHLMAG	471				
QY	457	ADEMLVSRFPFCGLIQLHAMHYGTVPVVA	STGGLVDVTKEGVTGFHMGALNPD--	KLDE	514		
Db	472	ADVLA	VSRFPFCGLIQLQGMRYGTGPCACASTGGLVDVTIEGKTGFHMGRLSDVCKVVEP	531			
QY	515	ADALAAATVRA	SEVAFAGGRYPENVANCITSQDLSWSPKQAKWEGLL-----	EEVYV	566		
Db	532	SDVKVATTLKRAIKIYGTITAYNEMVNCNMQLSWKGP	AKNWEINVLGLGVAGSEPGVE	591			
QY	567	KGKGVATAKKEIKVP	582				
Db	592	GE-ETAPLAKENVAAP	606				
RESULT	13						
Q9MAQ0							
ID	Q9MAQ0						
AC	Q9MAQ0						
		PRELIMINARY;	PRT;		610	AA.	

DT	01-OCT-2000 (TREMblrel. 15, Created)
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE	Granule-bound starch synthase.
DN	F9111.8.
GN	
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX	eurosid II; Brassicales; Brassicaceae; Arabidopsiis.
NCBI_TaxID=3702;	
[1]	
RN	SEQUENCE FROM N.A.
RC	STRAIN=Cv. COLOMBIA;
RA	Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA	Altafi H., Nguyen M., Lam B., Southwick A., Bei, Buehler E., Chin C.,
RA	Chlou J., Choi E., Dunn P., Gonzalez A., Howng B., Kim C., Koo T.,
RA	Lee J.M., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N.,
RA	Pham P., Sakano H., Schwartz J., Shinn P., Thaveri A., Toriumi M.,
RA	Vaysberg M., Walker M., Yu G., Ficker J., Theologis A., Davis R.W.;
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AC06424; AAF31273.1;
DR	InterPro; IPR001296; Glycos.transf_1.
DR	Pfam; PF00534; Glycos.transf_1; 1.
SQ	SEQUENCE 610 AA; 66879 MW; CF17F25BBL2220DF CRC64;
Query Match	38.6%; Score 1405.5; DB 10; Length 610;
Best Local Similarity	49.6%; Pred. No. 2.3e-74;
Matches 299;	Conservative 76; Mismatches 163; Indels 65; Gaps 14;
Qy	3 VASTSRPSSARPIVINAASFVKYKTANQLRELARGSARKTSFSAVTGATGC--ALD 60
Db	49 VDNLRQRSAKPV-----SAKSKRSKKVTAGKIYCEKGMS 85
Qy	61 IVMAAEVAPWPKSGGLGVDTGGPLIEIVKRGHVMITAPRYDOYADAWDTSVVVDI-MG 119
Db	86 VIFTGAEGVPWSKSGGLGDVLGGLPALAARGHRVMTICPRYDQYKDWDTCVVVQIKVG 145
Qy	120 EK----VRYFHISIKKVGRHVMTIDHPWFIAKWGKTGSKYLGPSPGADYLDNHKRFALFCKA 176
Db	146 DKVENVRFFHCYKRGVDRVFVDHFIFTLAKVYGKTSKIYGITGYVDNDNLRFSLCOA 205
Qy	177 ATEAAILPF-----GP-CEDCVFVANDWHSAIPLVLLKDEXQPGOFTKAKSVLAIH 228
Db	206 ALEAPQVNLNLSKYFSFGPYGEDVVFVANDWHETALLPCYLSMKQSRGVYNNAKVVFCLH 265
Qy	229 NATAQGRMWEBAFKDKLP---QAADFKLAFSDGYAKVYVEATEPMEEDEKPPLTKGTYYK 285
Db	266 NIAYQGRAFDYSILLNPISKFSSFDPMDGYEK-----PVKG---RK 305
Qy	286 INWLKGGIIAADKLVTVSPNATETAADAAGGVELDTVIIRAKIGIINGMDIEENPKT 345
Db	306 INWKAAILEAHRLVTVSPYAQELISGDRGVELHKYLRMTVSGIINGMDVQWNEST 365
Qy	346 DKFLSAPYDQNSVTAGKAARKEALQAEELGVDPDPAPLFAFTIGRLEEOKGVDIILAALPK 405
Db	366 DXYIDIKDYITVTDAPLKLIKEALOAAVGLPVDRPVPIGFIRLEEOKGSIDLVEAISK 425
Qy	406 ILATPKVOIAILGTCKAAEYKLVNAIGTKYKGRAGKVYKFSAPLAHMHTAGADFMVVSFR 465
Db	426 FMGL-NVQMVLITGTTKKWEAQILLEEFPGKAVGAKFNPNVAHMITTAGADFTIVPSR 484
Qy	466 FEPCGLIQLHAMHYGTVPVASTGCLVDTVKEGVTGFHMGALNP-DKILDEADADALAAT 523
Db	495 FEPCGLIQLHAMRYGTVPFIVASTGSLVDTVKDGYTGTFHLGRNVKCEVVDPDVDTATAKA 544
Qy	524 VRRASEVPAGGRRYPENVANCISQDLSSWSKPAQKWEGLLEEVYVYGGVATAKKEEIKVPV 583
Db	545 VTRAVAVYGTSAMQEWNKMDQDSFKWGPARLWEKVLSSLNVA-GSEAGTGEETI-APL 602
Qy	584 AEK 586
Db	I::
Db	603 AKK 605

This Page Blank (uspto)

Result No.	Query %			ID	Description
	Score	Match	Length		
1	1426	39.2	609	3	US-08-941-445A-7
2	1377.5	37.8	600	4	US-09-388-743-22
3	1358.5	37.3	533	3	US-08-941-445A-5
4	1358	37.5	616	4	US-09-388-743-14
5	1327.5	36.5	615	4	US-09-388-743-2
6	1316.5	36.2	614	4	US-09-388-743-18
7	903	24.8	641	4	US-08-836-567-10
8	885.5	24.3	649	4	US-09-192-909-2
9	884.5	24.0	601	4	US-09-196-390-2
10	875.5	24.0	801	4	US-09-388-743-26
11	873	24.0	767	4	US-08-836-567-8
12	869.5	23.9	583	3	US-08-941-445A-13
13	865.5	23.8	539	3	US-08-941-445A-21
14	858	23.6	690	4	US-09-388-743-6
15	847.5	23.3	698	3	US-08-941-445A-11
16	847	23.3	558	4	US-08-836-567-6
17	842	23.1	799	4	US-09-196-390-6
18	726	19.9	459	4	US-08-836-567-4
19	702.5	19.3	669	3	US-08-941-445A-9
20	580	15.9	477	1	US-07-735-065-2
21	580	15.9	477	1	US-08-469-202-12
22	580	15.9	477	2	US-08-484-434C-12
23	570.5	15.7	677	4	US-08-836-567-2
24	570.5	15.7	1197	4	US-08-836-567-12
25	570.5	15.7	1230	2	US-08-968-542C-35
26	559	15.4	735	4	US-09-115-704-2
27	534.5	14.7	533	4	US-09-388-743-10

QY 49 VTGATGATCALTDMVAAEAPVSKTGGTGLDVTGGLPIELVKRGHRYMTIAPRYDOYADA 108
Db 75 VVTAIGA--GNNVYVGAEMAPSKTGGTGLDVTGGLPPAMAANGHRVWVSPRYDOYADA 132
QY 109 WDTSVVVDI---MGEKRVYFHSIKKGVHRVWIDHPWFLAKVWGTGSKLYGPRSGADYL 164
Db 133 WDTSVVVAEIKVADRYVRFFHCYKRGVDRVFIIDHPSFLEKVMGTGKIYGPDTGVYDK 192
QY 165 DNHRKRALFCKAALEAARVL-----PRGP-----GEDCVFVANDWHSALVPVLLKDEYQPKG 216
Db 193 DNQMRFSILCOALEAPRILNINNPNYFKGTGYGDEVVFCNDWHTGTLASLKNYQPNQ 252
QY 217 QFTRAKSVLAHNIATFOGRMWEAEKDTKLPQAAFDKLAFLSDGYSAGYAKVYTEATPMEDEKP 276
Db 253 IYRNAKVAFCIHNSYQGRFAFEDYPELNLSERFSDFIDGY-----DTPVEG----- 302
QY 277 PLTGKTYKKINLWKGIIAADKLVTVSPNYATEIAADAAGVELDVTIRAKGIEGTNGM 336
Db 303 -----RKINMKAGILEADRVLTSPYAEELISGTARGCELDNIMRLTGTITGVNGM 355
QY 337 DIEWNPNTDFLSAPYDQNSVYAGKAAKAAKEALQAEGLPVDPTAPLFAFICRLLEEOKGV 396
Db 356 DVSEMDPSKDIYITAKYDATTALAKALNKEALQAEGLPVDKPIPLAFICRLEEOKGP 415
QY 397 DIIAALPKILATPKVQIAILGTGKAAYEKLVAIGTKYGRAGKGVWFSAPLAHMLTAG 456
Db 416 DVMRAAIIPELM--QEDVQIILVLTGKKFELKLSMEEKYPGKVRVAVKFNAPLAHMLTAG 474
QY 457 ADMVLVPSRFEPCGLIQLHAMHYGTVPVVAAGVSTGGLVDVTKGVTGFHMGALNPD--KLDE 514
Db 475 ADVLAVPSRFEPCGLIQLQGMRYGTGCACASTGGLVDVTKGVTGFHMGRLSVDCKVVEP 534
QY 515 ADADALATVRRASVFAAGRYPEMVANCISQDLSWSKPAQKWEGLLEEVYVKGKGG---- 570
Db 535 SDVKKVAATLRAIKVGTGTAYEVMVNCMDLSWKGPAKKNWNL--LGLGVAGSAPG 592
QY 571 -----VATAKKEEIKVP 582
Db 593 IEGDEIAPLAKENVAAP 609
RESULT 2
US-09-388-743-22
; Sequence 22, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Tulipa fosteriana
US-09-388-743-22
Query Match 37.8%; Score 1377.5; DB 4; Length 600;
Best Local Similarity 49.5%; Pred. No. 4.3e-112;
Matches 279; Conservative 79; Mismatches 165; Indels 41; Gaps 10;
QY 42 KTSR--SAVTGATGATCALTDMVAAEAPVSKTGGTGLDVTGGLPIELVKRGHRYMTIA 99
Db 54 KSTPKVGSVVAPTVETGMNLVFGTGTGYSKTKGGLDVLGGLPPALAARGHRYMVT 113
QY 100 PRYDOYADAWTSSVVVDI--MGEK--VRYFHSIKKGVHRVWIDHPWFLAKVWGTGSKLY 155
Db 114 PRYDOYADAWTSSVVVDI--MGEK--VRYFHSIKKGVHRVWIDHPWFLAKVWGTGSKLY 173

QY 156 GPRSGADYLDNHKRALFCKAALEAARVLPF-----GP--GEDCVFVANDWHSALVPVL 207
Db 174 GPVGTGTDYDNDQLRFSLLCOALEAPRVLNLSNNSKYFSGPYGEDVYFTIANDWHTGTLPCY 233
QY 208 LKDEYQPKGQFTKAKSVLAHNIATFOGRMWEAEKDTKLPQAAFDKLAFLSDGYSAGYAKVYTEA 267
Db 234 LKSVYKSGLYESAKVAFCIHNMAYQGRFAPDFSLNLPDTOTFKSSDFEFGYTK----- 288
QY 268 TPMEDEKPPITGTYKKINLWKGIIAADKLVTVSPNYATEIAADAAGVELDVTIRAK 327
Db 289 -----PVKG-----RKINMKAGILEADVTVPYAKELVSGDEGRGVDLNDVRLR 336
QY 328 GIEGIVNGMDIEENWPKTDKLSAPYDQNSVYAGKAAKAAKEALQAEGLPVDPTAPLFAFI 387
Db 337 GVKGVINGMDTNVWNPITDKFITANYDATMTVEAKRVNKQELQAEVGLPVDPIVIVFV 396
QY 388 GELEOKGVDIILAALPKILATPKVQIAILGTGKAAYEKLVAIGTKYGRAGKGVWVKTSA 447
Db 397 GRLEOKGSDIILAAIPELM--DENVQIILVLTGKKHLEKELEIEBEQPDKMLVAKFN 455
QY 448 PLAHMLTAGADFMVPSRFEPCGLIQLHAMHYGTVPVVAAGVSTGGLVDVTKGVTGFHMG 506
Db 456 PLAHMMAGSDIILAAIPELM--DENVQIILVLTGKKHLEKELEIEBEQPDKMLVAKFN 515
QY 507 -LNPDKLEADADALAAVVRASVFAAGRYPEMVANCISQDLSWSKPAQKWEGLL---- 561
Db 516 TVECETVDPVDVAGIVTKVRAKLYGTGTPAFSEMVCNMAODHWSKPEAKKWEILLGLE 575
QY 562 ---EEVVYKGVGAVATAKKEEIKVP 582
Db 576 VDGSEFGDGEIAPLAKENVAAP 599
RESULT 3
US-08-941-445A-5
; Sequence 5, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Haining
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid

	QY	67	EVA	PWS	TGGLGDV	TGGLPIELVKR	GHRWVTIAPRYDQYADAWDTSVVVVDI-MG---	BKV	122
					:				
					:				
	dB	97	EVA	PWS	TGGLGDV	LGPLPMAAK	GHRWVTIAPRHQDYKDGWDIAVFVLKVGDRITV		156

123 RYRHSIRKGVHRVWIDHPWFLAKVWGKTGSLYGRSGADYLDNHRKFALECKAAIEAAR 182
 157 REFHCYKGVDRVDFVDRHFLFLEKVGKGGKIYGPVITDVEDNOLRCLCLATLETPR 216
 183 VL-----PFGP-GEDEVFVANDHWSALVPLVKD-EYQPKGQFTKAKSVIAIHNIAPQ 233
 217 VLPNNNNYHSGPKGED-LFIANDWHTALLPCYLKTIYQAHGIYKNAKVAFCIHNIAYQ 275
 234 GRMEEAPKDKLQOAAFDKLAFLSDGYAKVYTEATPMEDEKPLTGKTKKINWLGKI 293
 276 GRFAFEDPSRLNLPDTRKSSDFIDGYAK-----PIKG---RKINWMAKGI 318
 294 IAADKLTVSPNYAIEAADAAGGVVELDTVIRAKGIEGVINGMDEEWNPKDKFELSAPY 353
 319 IESDRALTVSPYIAQELVSGIDKGVLELDNIIRLKTICGLINGMDINENWNSDKITANY 378
 354 DONSYYAGAAKAAEQALQELGPDPTAPLFAFTIGRLEEQKGVDIILALPKILATPKVO 413
 379 DATTVMEAKPLNKEALQAEVLGPNVKIPVIAFAGLEEQKGSIDILAEAIKPF--DODVQ 437
 414 TAILGTGAAYEKLVAITGKYKGRAGVVKESAPLAHMLTAGADFMVPSRFEPCGLIQ 473
 438 VIVLTGTRKKLRQALLEDPEPDKFRAHMKFNIPLAGIMAGADILVTPSRFEPCGLIQ 497
 474 LHAMHYGVPPVASTVGGVLDVTKVKGVTGFHMG--ALNPDKLDEADADALAAATVRRASEVF 531
 498 LOGMYGTPSMCTITGGVLDVTKVKGITGFHMGPFSEVCEADIADADVILKIVEAKRALMY 557
 532 AGGPEYPMVANCISQDLSSKSPAKWEGLL-----EEVYKGGGVATAKKEEKVP 582
 558 GTPAFEMIQNMAQDFSKGPAKEWKEKFLSLGLESGEAGIEGE-EVAPLAKENVATP 615

RESULT 6

US-09-388-743-18
 ; Sequence 18, Application US/09388743
 ; Patent No. 6423886
 ; GENERAL INFORMATION:
 ; APPLICANT: Singleary, George
 ; APPLICANT: Zhou, Lan
 ; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
 ; TITLE OF INVENTION: Use in the Production of New Starches
 ; FILE REFERENCE: 1144
 ; CURRENT APPLICATION NUMBER: US/09/388,743
 ; CURRENT FILING DATE: 1999-09-02
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 18
 ; LENGTH: 614
 ; TYPE: PRT
 ; ORGANISM: Typha latifolia
 ; US-09-388-743-18

Query Match 36.2%; Score 1316.5; DB 4; Length 614;
 Best Local Similarity 47.2%; Pred. No. 9.8e-107;
 Matches 271; Conservative 85; Mismatches 165; Indels 53; Gaps 13;
 42 KTSRSRAVTGATGAT-----C--ALDIVMVAEAPWSTKGTGGVGTGGIPIELVK 90
 60 KTSKQFKNGCSTVNLPRPRAVYICGKGMNLVFGAEMAPWSTKGTGGVGLGGLPPLAA 119
 91 RGRVMTIAPRYDQADAWDTSVVVDI-MG---EKVYFHSIKKGVHRVWIDHPFLAKV 146
 120 NGRVMVIAFRYDQYMDAWDTDALVELKVGDRCTVREFFHCYKRGVDRVDFVHPMFLAKV 179
 147 WGTGKSLYGRSGADYLDNHRKFALECKAAIEAARVLPF-----GP-GEDEVFVAND 198
 180 WGTGKGIYGTGTGTDQDNOLREFSLCQAALEAPRILNLSNDSFGSPGIEDVIFIND 239
 199 WHSALVPLVKDEYQPKGQFTKAKSVIAIHNIAPQGRMWEAEAFKDTKLPOAFAFDKAFSD 258
 240 WHISLLPCYLKSMVHPRGYKNAKVAFCIHNISYQGRFSPSPDFEFLNLPENFKSSFSFID 299

259 GYAKVYTEATPMEDEKPLTGKTKKINWLGKIADKLTVSPNYATEIAADAAGGV 318
 300 GYK-----PVKG---MKINWMAKGIKESDRVFTVSPYTAQELLSGEERG 342
 319 ELDTVIRAKGIEGVINGMDIEBWNPKDKFISAPYDQNSVYAGKAAKAAEQALQELGPDV 378
 343 ELDNILRTVTSITGVINGMDVNEWNPLTDKYISVNDYAKPVMEAKPLNKEALQAESWLACR 402
 379 PTAPLFAFTIGRLEEQKGVDIILALPKILATPKVQIATLGTGAAYEKLVAITGKTKK 438
 403 QSHPCNCHIRLEEQKGSIDILAAIPEIM-DENVOLILGTGKEMENQLESMEEMFPDK 461
 439 AGVYKFSAPLAHMLTAGADFMVPSRFEPCGLIQLHAMHYGVTPVVAFTGGVLDVTKV 498
 462 VRAMKFNAPLAHOMTAGADIIVTPSRFEPCGLIQLQGMGYGTPSPACSSTGGVLDVTKV 521
 499 VTGFHMGALNP--DKLDEADADALAAATVRRASEVFAGGYPPEMVAANCISQDLSSKSPAK 556
 522 KTFHMGPFSAECEVVDLSKVYTVTKRALKVICTPAFEDMKNKMAQDLSSKSPAKN 581
 557 WEGLEEVYKGGVATAKKE-----EEIKVPVPAEK 586
 582 WEQVLLNL-----GVAGSEPGYDGEI-VPLAKE 609

RESULT 7

US-08-836-567-10
 ; Sequence 10, Application US/08836567
 ; Patent No. 6130367
 ; GENERAL INFORMATION:
 ; APPLICANT: Kossmann, Jens
 ; APPLICANT: Springer, Franziska
 ; APPLICANT: Abel, Gernot
 ; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
 ; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
 ; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FISH & NEAVE
 ; STREET: 1251 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10020
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/836,567
 ; FILING DATE: 24-JUL-1997
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/EP95/04415
 ; FILING DATE: 09-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DE P 44 41 408.0
 ; FILING DATE: 10-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haley Jr., James F.
 ; REGISTRATION NUMBER: 27,794
 ; REFERENCE/DOCKET NUMBER: Agrevo-4
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-596-9000
 ; TELEFAX: 212-596-9090
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 641 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-836-567-10

```

Query Match      24.8%; Score 903; DB 4; Length 641;
Best Local Similarity 38.1%; Pred. No. 1.7e-70;
Matches 219; Conservative 91; Mismatches 183; Indels 82; Gaps 20;

QY 30 QLLRELARGSARKSTSRSAVTG-----ATGAFCAIDIVMVAEVPWSKTGG 76
Db 89 QLIPHSVAGDATVESHVIDVANDRDLDSETERMEETPIKLTNIFVTAAAPYSKTGG 148
QY 77 LGDVTGGLPELVKGRHVTIAPRY-----DQYADAWDTSV--VVDIMG--EKVRYF 125
Db 149 LGDVCGLFMAAARGHRVWVSRYLNGGSPDEKYANAVDLDRATVCFEGDAQEAVFY 208
QY 126 HSIKKGVRHWIDHPFLAKVWGKSKLYGPRSGADYLDNKHKRFALFCKAAEAARVLP 185
Db 209 HEYRAGVDVVDHSSYC-----RPGTP-YGDIYGA-FGDNQFRFTLLSHAACEAPLVLP 261
QY 186 FGP---GECVVFVANDHWSALVPVLKDEYQPKGQFTKAKSVLAIHNIHAFQGRMWEAEK 242
Db 262 LGGFTYGEKCLFLANDHWAALVPLLAAKYRPGYGVYKARSIVAIIHIAHQVEPAVTYN 321
QY 243 DTKLPOAAFDKLAESDGYAKVTEATPMEDEKPPLTGTYKKINLWKGIIIAADKLVTY 302
Db 322 NLGLPQWGA---EWIPTWARAHALD-----TGCT---VNVKGAIAVADRILTV 368
QY 303 SPNYATEAADAAGVEIDTVI---RAKGTEGIVNGMDIENWPKDKFLSAPYDONSVA 360
Db 369 SOGYSWEITT-PEGYGLHELLSSRSQVLNGITNGIDVNDWNPSTDEHTASHYSINDL-S 426
QY 361 GAAAKEALQALQELGPVDPFAPFAGIRLEBQKVDIILAALPKILATPKVOIALILGTG 420
Db 427 GKVOCKTDLQELGUPIRPDCPLGIFGLRDKYQKGVDIILSAIPELMQN-DVQVVMGSG 485
QY 421 KAAVEKLVNAICTYKGRAGKGVKFSAPLAHMLTAGADFMVLPSPRFEPCGLIQLHAMHYG 480
Db 486 EKQVEDWNRHTENLFKDFRAWGVNVPVSHRITAGCDILLMPSRFEPCGLNQLYAMRYG 545
QY 481 TVPVVASTGGVLVDIVK-----EGV---TGFMHGMALNPKDIDEADALAAATVRASEV 530
Db 546 TIPIVHSTGGLDVTKDNPVYAGIEGIGSGTGTWTFSLTSEKL-----LDTLKLA 595
QY 531 FAGGRYPE-----MVANCISQDLSWSKPAQKWE 558
Db 596 ---GTYTEHKSSWEGLMRRGMGRDYSWENAAIQYE 627

RESULT 8
US-09-192-909-2
; Sequence 2, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Froberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; TITLE OF INVENTION: starch synthases from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: PCT/EP97/02527
; FILING DATE: 16-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-556-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-192-909-2

```

```

Query Match      24.3%; Score 885.5; DB 4; Length 649;
Best Local Similarity 37.8%; Pred. No. 6.1e-69;
Matches 207; Conservative 97; Mismatches 178; Indels 65; Gaps 16;

QY 42 KTSRSVATGATGATCALDIVMVAEVPWSKTGGGLDVTGGPLPIELVKRGHVRVMTIAPR 101
Db 132 KEQARAKVT-----QNVFVTGEASPYAKSGGLGCVGSLPVALAARGHRVWVMPR 183
QY 102 Y-----DQYADAWDTSVVDIM---GE-KVRYFHSIKKGVRHWIDHPFLAKVWGK 151
Db 184 YLNGTSDKNYANAFYTEKHIRIPCFGEHEVTFHEYRDSVDVWVVDHPSY-----HRP 237
QY 152 SKLYPRSGADYLDNKHKRFALFCKAAEAARVLPFP---GEDCVFVANDHWSALVPVLL 208
Db 238 GNLVGDKFGA-FGDNQFRYTLICYAACEAPLVLELGGYIYGNCMFVVDHWSALVPVLL 296
QY 209 KDEYQPKGQFTKAKSVLAIHNIHAFQGRMWEAEFKDTKLPOAAFDKLAESDGYAKVTEAT 268
Db 297 AAKYRPGYGVKDSRSILVHNLAHOGVEPASTYPLGLPPEWYGALEW-----VPEWA 350
QY 269 PMEEDKEPPLGTYKKINLWKGIIIAADKLVTVSPNYATEIAADAAGVEIDTVI--RA 326
Db 351 RRHALDKG-----EAVNFKLGAVTADRIVTVSKGYSWEVTT-AEGGQGLNELLSRK 402
QY 327 KGIIEGIVNGMDIENWPKDKFLSAPYDONSVAAGKAAKEALQALQELGLPVDPTAPLRAF 386
Db 403 SVLNGIVNGIINDNPAIDKICPHYSVDDL-SGAKCKGALOKELGPIRPDVPVLIGF 461
QY 387 IGRLEEQKGVDIILAALPKILATPKVQITAILGTGKAAVEKLVNAICTYKGRAGKGVKFS 446
Db 462 IGRLDYQKGLDILIIIPDLN-REDVQFVMLGSGDPELEDNMRSTESIFKDKFRGWGFS 520
QY 447 APLAHLMTAGADFMVLPSPRFEPCGLIQLHAMHYGTVPVVASTGGVLVDIVK-----E 497
Db 521 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGVTPVWVHATGLRDTVENFNGE 580
QY 498 GVTGFHGMALNPKDIDEADALAAATVRASEVAGGR--YPMVANCISQDLSWSKPAQ 555
Db 581 OGTGWAFAPLTTN-----MLWTILRTALSTYREHKSSWEGLMKRGMSKDFTDWHAEE 632
QY 556 KWEGGLE 562
Db 633 QYEQIFQ 639

```

```

RESULT 9
US-09-196-390-2
; Sequence 2, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie

```

```

RESULT 10
US-09-388-743-26
; Sequence 26, Application US/09388743
; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Typha latifolia
; US-09-388-743-26

```

Query Match	24.0%	Score	875.5	DB 4	Length	801
Best Local Similarity	36.5%	Pred. No.	6.4e-68			
Matches	225	Conservative	85	Mismatches	199	Indels
	107	Gaps				
QY	2	AVASTSRPSS-----	ASRPVINAASFGVKKTANQLLR	ALARGSARKSTSR	-----	47
Db	156	AVDAFGPEPRLGTTKILSPFYLEASDGAKE	NAEDLVEAKLDSVHVVDKDL	PGPEENE	215	
QY	48	---AVTCATGATCATLDIVKAAAEVAPSK	TGGIGDVTGGLP	PIELVKRGHRVMT	LAPRYDQ	104
Db	216	VPILAGAN----	VNNIIIVAAECAPSKTGGSLGDVAGALP	KPALARRGHRVMMV	VAPRYGN	271
QY	105	YADAWDTSV--VVDIMGE--	KVRVFHSIKKGHRVWIDH	PFLAKVWGKTSKLY	-GPRS	159
Db	272	YAEQDIDGVRKYKVVHVGDMETVYF	HAIDGVDFVMDSPDFR	-----	GNRIYEGNR	325
QY	160	GADYLDNKHRELFALCKAIAEARVLP	-----	FGPGEDCVFV	VANOWHSLVPLIKDEYQP	214
Db	326	---VDILKRMILFCKAAVEVPWHV	PCGFCYGDG	-NLAFITNDWHTALLP	PYLKAYYRD	380
QY	215	KGQFTKAKSVLAHNIAFQGRMWEAE	FVDTKLPQAADF	KLAFSDGAKVYVTEATP	MEISDE	274

```

Db 381 NGLMKYARSLVITHNIAHQGRGPDVDFKFGVGLPDHYLDLFLYD----- 424
QY 275 KPPLTGTYYKINLWKGIIAADKLVTYSPNYATEIAADRAAGVVELDTVIRAKG--IEGI 332
Db 425 --PVGG---EHLNFAAGLTADRVTVSHGYAWELKT-SEGGMGLHEIINESNWKFGI 478
QY 333 VNGMDIEWPKTKDLSAP-----YDONSVYAGKAAKALQAEGLPVPDPTAPLEAFIG 388
Db 479 VNGIDAKESPEFDVHLKSDGYTNSLDTLEMGKPCVKCAALQREVGLPVRDNPVPIAFIG 538
QY 389 RLBEOKGVDIILALPKILATPKVQIAILGTGKAAAEKLVNAIGTKYKGRAGKGVKFSAP 448
Db 539 RLHQKGVDIILAEAMPWIV-SHDVQVVMGLTGTRODLENNLRNEEGHRDKVRAWAFSVK 597
QY 449 LAHLTAGADFMVPSRFPCGLIQLHAMHYGTVPVYASTGGVLVDIVKEGVTGHEHMGALN 508
Db 598 MAHRIITAGADILMPSEFEPCGLNQLYAMMIGTIPVVHAVGGLRDTVTO----- 646
QY 509 PKLDEADADALATVRRASEVFAGGRYPPEMVA NCIS-----QDLSW 550
Db 647 ---FDPFNESSGLGWTFDRAE---AGKLIHALNNCLNTYWNKYDSWKGLQTRGMQDLSW 699
QY 551 SKPAQKWEGLLEEVY 566
Db 700 DNAAQYEDVLVAKY 715

RESULT 11
US-08-836-567-8
; Sequence 8, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836.567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
US-08-836-567-8

```

```

Query Match 24.0%; Score 873; DB 4; Length 767;
Best Local Similarity 39.7%; Pred. NO. 9.9e-68;
Matches 211; Conservative 78; Mismatches 183; Indels 60; Gaps 14;

```

```

QY 56 TCALDIVMAAEVAPWSKGTGLGVDVGTGELVKRGRHVMVTIAPRYDQYADADWTSV-- 113
Db 273 TNVNNIILVASECAPWSKGTGLGVDVAGALPKALARRHVRVMVAPRYDVPQDSGYRK 332
QY 114 --VVDIMGEKVRFHSIKKGVHRVIDHPFLAKVWKGTSKLYCPRSGADYLDNHRFA 171
Db 333 IYKVGQDVETTFQAFIDGVDFID-----SHMERHIGNNIY----GGNRVDILRMV 383
QY 172 LFCKAAATEAARVLP-----FGPGEDCVFVANDHSHALVPVLLKDEYQPKGQFTAKSVLA 226
Db 384 LFCKAAIEVPHWPCGGVGYGDG--NLVFIANDNHTALLPVYLKAYYRDGIMNYTRSLV 442
QY 227 IHNIAFGRMWEAFKDKLPQAAFDKLAFSGKYAKVYTEATPMEDEKPLTGKTYKKI 286
Db 443 IHNIAHQGRGLEDFSYVDLPPHYMDPKLYD-----PVGGEHP--- 481
QY 287 NMLKGLIADKLVTYSPNYATEIADAAGVELDTVIRAKG--IEGVNGMDIEWNPK 344
Db 482 NIFAAGLKTADRVTVSHGYSWELKT--SQGGWGLHQIINENDWKLOGIVNGIDTKEWNP 540
QY 345 TDKFLSAP-----YDONSVYAGKAAKALQAEGLPVPDPTAPLEAFIGRLEOKGVDIIL 400
Db 541 LDVHLQSDGYMNSLDTLQTKPQCKAAALQKELGSLPVRDVPVLTGTRLDPCQGVDLIA 600
QY 401 AALPKILATPKVQIAILGTGKAAAEKLVNAIGTKYKGRAGKGVVYKFSAPLAHMLTAGADM 460
Db 601 EASAWMMG--QDVQLVMLGTGRDLEQMLROFECQHNNDKINGWVGFVSKTSHRITAGADIL 659
QY 461 LVPSRPEPCGLIQLHAMHYGTVPVYASTGGVLDTVK-----EGVTGFHMGALNPDKLDE 514
Db 660 LMPSRPEPCGLNQLYAMKYGTIPVVHAVGGLRDTVQFPDFNPSGLGWTFSRAESQLIH 719
QY 515 ADADALAAATVRRASEVFAGGRYPPEMVA NCISQDLSWSKPAQKWEGLLEEVY 566
Db 720 ALGNCL-LTYREYKSWEG-----IQTRCMTQDLSWDNAAQNYEEVLIAKY 765

```

```

RESULT 12
US-08-941-445A-13
; Sequence 13, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941.445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:

```

COUNTRY: US
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/941.445A
 FILING DATE: 30-SEP-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026.855
 FILING DATE: 30-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Winner, Ellen P
 REGISTRATION NUMBER: 28,547
 REFERENCE/DOCKET NUMBER: 89-97
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 539 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-941-445A-21

Query Match 23.8%; Score 865.5; DB 3; Length 539;
 Best Local Similarity 40.0%; Pred. No. 2.6e-67;
 Matches 198; Conservative 84; Mismatches 158; Indels 55; Gaps 14;

QY 42 KTSRSVATGATGATCALDIVMAEVAEAPSWKGGDVTGGLPIELVKRGHVTIAPR 101
 DB 40 KEQARAKVTQS-----IVFVTGEASPYAKSGGLDVGCSLPVLAARGHVRVVMVMPR 91
 QY 102 Y-----DOYADAWTSVVVDIM--GE-KVRYFHSIKKGVHVRVMDHPFLAKVWGKTG 151
 DB 92 YLNGTSDKNYANAFYTEKHIRIPCFGEHEVIFFEYRSDVWVFDHPSY-----HRP 145
 QY 152 SKLYGPRGADYLDNHRKFAFCAAIEAARVLPFGP---GEDCVFVANDHWSALVPVLL 208
 DB 146 GNLGDKFGA-FGDNQFRYTLCCYAAACEAPLILELGGYIYGQNCMEFVNDHWSALVPVLL 204
 QY 209 KBYQKQGTAKSVLAHNTAFQGRMWEAEFKDKLPQAAFDKLAFAFDGYAKVYTEAT 268
 DB 205 AAKYRPGYVYKDSRSILVHNLAHQGVPEASTYPDGLGPPPEWYGALEW-----VPEWA 258
 QY 269 PNEDEKPPLTGTYKKINLKGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RA 326
 DB 259 RRHALDKG-----EAVNFKGAVVTADRIYTVSKYSWEVTT-AEGGGLNELLSRK 310
 QY 327 KGIEGVNGMDIEWNPKTDFLSAPYDQNSVYAGKAAAEALQAEGLGLPVDPTAPLFAF 386
 DB 311 SVLNGVINGIDINDWNPATDKCIPCHYSVDLL-SGAKCKGALQKELGPIRDPVPLIGF 369
 QY 387 IGRLEQGVDTIILAAALPKILATPKVOIALGTGKAAAYEKLIVNAIGTKYGRAGVYKFS 446
 DB 370 IGRLDYQKGDILQIILIPDLN-REDVQFVMLGSGDPELDWMRSTESIFKDKFRGWGFS 428
 QY 447 APLAHLMTAGADMVPSRFEPCGLIOLHAMHYGTVPVVAATSGGLVDTVK-----E 497
 DB 429 VPVSHRITAGCDILLMPSRFEPCGLNQLYAMQYGTVPVVAATSGGLRDTVENFNGFGE 488
 QY 498 GVTGFHMGALNPOKL 512
 DB 489 QGTGWAFAPLTENN 503

RESULT 14
 US-09-388-743-6
 ; Sequence 6, Application US/09388743

NAME: Winner, Ellen P
 REGISTRATION NUMBER: 28,547
 REFERENCE/DOCKET NUMBER: 89-97
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 583 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-941-445A-13

Query Match 23.9%; Score 869.5; DB 3; Length 583;
 Best Local Similarity 38.3%; Pred. No. 1.3e-67;
 Matches 209; Conservative 90; Mismatches 174; Indels 73; Gaps 17;

QY 6 TSPSSARPIVINA-----ASFQVYKKT---ANQLRELARGSARKTSRSVAT 50
 DB 36 TGEFASPTPPPDAGLDGLLEPEGIAEGSIDNTVVVASEQDSEIVVG---KEQARAKVT 92
 QY 51 GATGATCALDIVMAEVAEAPSWKGGDVTGGLPIELVKRGHVTIAPR-----DQ 104
 DB 93 QS-----IVFVTGEASPYAKSGGLDVGCSLPVLAARGHVRVVMVMPRYLNGTSDKN 144
 QY 105 YADAWTSVVVDIM--GE-KVRYFHSIKKGVHVRVMDHPFLAKVWGKTGSKLYGPRSG 160
 DB 145 YANAFYTEKHIRIPCFGEHEVIFFEYRSDVWVFDHPSY-----HRPGLNGDKFG 198
 QY 161 ADYLDNHRKFAFCAAIEAARVLPFGP---GEDCVFVANDHWSALVPVLLKDEYQPKQ 217
 DB 199 A-FGDNQFRYTLCCYAAACEAPLILELGGYIYGQNCMEFVNDHWSALVPVLLAAKYRPGV 257
 QY 218 FTKAKSVLAHNTAFQGRMWEAEFKDKLPQAAFDKLAFAFDGYAKVYTEATPNEDEKPP 277
 DB 258 YKDSRSILVHNLAHQGVPEASTYPDGLGPPPEWYGALEW-----VPEWARHALDKG- 310
 QY 278 LTKYKINLKGIIAADKLVTSPNYATEIAADAAGGVDELDTVI--RAKGIEGVNG 335
 DB 311 -----EAVNFKGAVVTADRIYTVSKYSWEVTT-AEGGGLNELLSRKSVLNGVING 363
 QY 336 MDIEWNPKTDFLSAPYDQNSVYAGKAAAEALQAEGLGLPVDPTAPLFAFGRLEEQK 395
 DB 364 IDINDWNPATDKCIPCHYSVDLL-SGAKCKGALQKELGPIRDPVPLIGFGRLDYQK 422
 QY 396 VDILILAAALPKILATPKVOIALGTGKAAAYEKLIVNAIGTKYGRAGVYKFSAPLAHMLTA 455
 DB 423 IDLIQIILIPDLN-REDVQFVMLGSGDPELDWMRSTESIFKDKFRGWGFSVPVSHRITA 481
 QY 456 GADFMVPSRFEPCGLIOLHAMHYGTVPVVAATSGGLVDTVK-----EGVTFHMG 506
 DB 482 GCDILLMPSRFEPCGLNQLYAMQYGTVPVVAATSGGLRDTVENFNGFGEQGTGWAFAP 541
 QY 507 LNPDKL 512
 DB 542 LITENN 547

RESULT 13
 US-08-941-445A-21
 ; Sequence 21, Application US/08941445A
 ; Patent No. 6107060
 ; GENERAL INFORMATION:
 ; APPLICANT: Keeling, Peter
 ; APPLICANT: Guan, Hanping
 ; TITLE OF INVENTION: Starch Encapsulation
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle
 ; CITY: Boulder
 ; STATE: CO

```

; Patent No. 6423886
; GENERAL INFORMATION:
; APPLICANT: Singletary, George
; APPLICANT: Zhou, Lan
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Curcuma zedoaria
; US-09-388-743-6

```

[illegible]

RESULT 15
 US-08-941-445A-11
 ; Sequence 11, Application US/08941445A
 ; Patent No. 6107060
 ; GENERAL INFORMATION:
 ; APPLICANT: Keeling, Peter
 ; APPLICANT: Guan, Hanping
 ; TITLE OF INVENTION: Starch Encapsulation
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle
 ; CITY: Boulder
 ; STATE: CO
 ; COUNTRY: US

ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/941.445A
 FILING DATE: 30-SEP-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026.855
 FILING DATE: 30-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Winner, Ellen P
 REGISTRATION NUMBER: 28,547
 REFERENCE/DOCKET NUMBER: 89-97
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8080
 TELEFAX: (303) 499-8089
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 698 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PS-08-941-445A-11

Query Match	23.3%	Score 847.5;	DB 3;	Length 698;
Best local Similarity	37.5%;	Pred. No. 1.5e-65;		
Matches 226;	Conservative	79;	Mismatches 204;	Indels 93; Gaps 19;

QY	4	ASTSRP----	SSARPIVINAAS	CGVKKTANQLL	RELARGSARKSTS	RSASVT-----	50		
		:	:	:	:	:			
Db	149	ASAVKPEP	GDDARPV----	ESIGIAE	PVD-----	AKADAAPAT	DAASAPYDREDNEP 198		
QY	51	GATGATC	ALDIVMAAE	VAEPWSKT	GLGDVDTG	GLPIELVK	RGRHVMTIAPRYDQYADAWD 110		
		:	:	:	:	:			
Db	199	GPLAGP	NMVMVVV	VASECA	PFCKTGG	GLGDVVG	ALPALARRGHRVMVVPYRGEYAEARD 258		
QY	111	TSV----	VVDIMEK	VRYPHS	IKKGVRH	WIDHPW	LAKVMGKTGSKLYGPRSGADYLDN 166		
		:	:	:	:	:			
Db	259	LGVR	RRYK	VAGQ	SEVITY	HSYIDG	VDVFVEAPPPFHR-----HNNIY-----GGERLDI 309		
QY	167	HKRFAL	CEKAAI	EAARV	LP-----	FGPEGD	CEVANDHSALVPVLLKDEYQPKGQFTKA 221		
		:	:	:	:	:			
Db	310	LKRMIL	CEKAAE	VPWPV	PCGGIV	YGDG--	NLVPFIANDWHITALLPVYIKAYYRDNGLMQYTA 368		
QY	222	KSVLA	THNTAF	QGRMEE	EAFKDT	KLPQAA	FDKLAFSDGYAKVYVTEATPMEDEKPLLTGK 281		
		:	:	:	:	:			
Db	369	RSVLV	THNTAH	OGRG	PVDDVF	DFDLPE	HYIDHFKLYDNI	GGDHS-----412	
QY	282	TYKIN	NLWKG	IIA	DKVLT	SPNATE	ITAADAAGVELDVTYRAG--IEGIVNGMDIE 339		
		:	:	:	:	:			
Db	413	-----	NVFA	AGLKT	ADRVV	TVSNGY	WELKT--SEGWGLHDIINQNDWKLQGI	VAGIDMS 466	
QY	340	ENNP	KTDK	FLSAP----	YDONS	VYAG	KAAEALQALGLPVDPTAPLFAFTRGLEBOGK 395		
		:	:	:	:	:			
Db	467	ENNP	AVDVLH	SDSYT	NTYF	ETLDT	GKQCKAALQRLQGLQYRDDVPLIGFCIRL	DHOGK 526	
QY	396	VDIIL	AALPK	ILAT	PKVQI	AILGT	CKAAEKL	VNAIGTKYGRAGKVYKFSAPLAHMLTA 455	
		:	:	:	:	:			
Db	527	VDIIL	ADAH	WHI--	AGQDV	OLV	MLGTGR	ADLEDMLRFESEHSDKVRAWGVFSVPLAHRITA 585	
QY	456	GADF	MLVPS	RFP	PCGLI	OLH	AMHYGT	VPVASTGGLVDITVKE---GYTGEMHGALNPDOK 511	
		:	:	:	:	:			
Db	586	GAD	ILLMP	SF	PCGL	NQLY	AMAYGT	VPVIRVAGGRDTPVAFDDPFDNDGLGW-----T 639	
QY	512	LDEAD	-----	DALA--	--	ATVRR	ASEV	FAGGRYP	EMVANCISODLSWSKPAQKWEGLLPEV 564
		:	:	:	:	:			
Db	640	FDRA	FNRM	IDL	ASH	LCH	LT	TYRNYK	ESWRACR-----ARGMAEDLSWDHAAVLYEDVLKA 694
QY	565	YY	566						

us-09-980-771a-3.ra1

Tue Jun 10 09:03:09 2003

Db 695 KX 696

Search completed: June 4, 2003, 14:58:30
Job time : 18.5962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:56:21 ; Search time 21.8651 Seconds
(without alignments)
3342.965 Million cell updates/sec

Title: US-09-980-771a-3

Perfect score: 3641

Sequence: 1 MAVASTSRSSARPVINA.....SASTSAAKPLVSAATRKSA 708

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1467.5	40.3	636	9 US-10-138-075-4	Sequence 4, Appli
2	1449.5	39.8	609	9 US-10-138-075-2	Sequence 2, Appli
3	1388.5	38.1	599	9 US-10-138-075-5	Sequence 5, Appli
4	884.5	24.3	671	10 US-09-952-677-2	Sequence 2, Appli
5	842	23.1	799	10 US-09-952-677-6	Sequence 6, Appli
6	562.5	15.4	915	9 US-10-163-214-6	Sequence 6, Appli
7	559	15.4	909	9 US-10-163-214-2	Sequence 2, Appli
8	555.5	15.3	914	9 US-10-163-214-12	Sequence 12, Appli
9	547.5	15.0	874	9 US-10-163-214-13	Sequence 13, Appli
10	410.5	11.3	474	12 US-10-007-693-99	Sequence 99, Appli
11	402.5	11.1	476	12 US-10-007-693-69	Sequence 69, Appli
12	385.5	10.6	293	9 US-10-163-214-10	Sequence 10, Appli
13	215.5	5.9	117	10 US-09-739-438-2	Sequence 2, Appli
14	206	5.7	409	9 US-09-738-626-4738	Sequence 4738, Ap
15	189	5.2	381	9 US-09-738-626-5896	Sequence 5896, Ap
16	164	4.5	143	10 US-09-739-438-4	Sequence 4, Appli
17	155.5	4.3	191	10 US-09-924-358-22	Sequence 22, Appli
18	155.5	4.3	418	9 US-09-738-626-3951	Sequence 3951, Ap
19	143	3.9	1222	10 US-09-137-531-15	Sequence 15, Appli

20	143	3.9	1252	10 US-09-137-531-9	Sequence 9, Appli
21	136.5	3.7	385	10 US-09-767-041-18	Sequence 18, Appli
22	135	3.7	2993	9 US-09-738-626-6239	Sequence 6239, Ap
23	132	3.6	2910	9 US-10-124-800-2	Sequence 2, Appli
24	130	3.6	1610	9 US-10-100-957A-22	Sequence 22, Appli
25	129	3.5	1125	9 US-10-100-957A-152	Sequence 152, Appli
26	128	3.5	3241	10 US-09-841-786-1	Sequence 1, Appli
27	126	3.5	387	9 US-09-738-626-3890	Sequence 3890, Ap
28	125.5	3.4	398	10 US-09-934-899-16	Sequence 16, Appli
29	125.5	3.4	398	10 US-09-934-868-36	Sequence 36, Appli
30	124.5	3.4	1704	9 US-09-991-262-40	Sequence 40, Appli
31	124	3.4	379	9 US-09-738-626-6655	Sequence 6655, Ap
32	124	3.4	1461	9 US-10-021-955-88	Sequence 88, Appli
33	123.5	3.4	413	9 US-09-738-626-4025	Sequence 4025, Ap
34	123	3.4	1528	9 US-09-945-917-3	Sequence 3, Appli
35	123	3.4	1583	9 US-09-945-917-4	Sequence 116, Appli
36	122.5	3.4	2756	9 US-10-331-061-7	Sequence 5590, Ap
37	122.5	3.4	3816	9 US-09-808-880-3	Sequence 5816, Ap
38	121	3.3	339	9 US-09-847-208-116	Sequence 12967, A
39	121	3.3	636	9 US-09-738-626-5590	Sequence 5277, Ap
40	120	3.3	2478	10 US-09-815-242-5816	Sequence 12605, A
41	120	3.3	2478	10 US-09-815-242-12967	Sequence 3810, Ap
42	118.5	3.3	372	10 US-09-815-242-5277	Sequence 2, Appli
43	118.5	3.3	384	10 US-09-815-242-12605	
44	118.5	3.3	795	9 US-09-738-626-3810	
45	118	3.2	267	10 US-09-852-555-2	

ALIGNMENTS

RESULT 1

US-10-138-075-4
; Sequence 4, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10/138,075
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Glycine max
US-10-138-075-4

Query Match	40.3%	Score 1467.5;	DB 9;	Length 636;
Best Local Similarity	51.0%	Pred. No. 1.4e-95;		
Matches	298;	Conservative 76;	Mismatches 159;	Indels 51;
Gaps	11;			
QY	24	VKTNQLRLRLARSARKSTSRSAVTGATGATCALDIYVAAEAPWSKGTGLGDTGG	83	
Db	79	IKTNVAQMRKPGOGKNRPKGM-----ITCGMTFIIIGTEVAPWCCKTGLGDLVGG	130	
QY	84	LPFELVKRGHRVMTIAPRYDOYADAWDSVVDI-----MGEKRVFHSIKKGVHRWIDH	139	
Db	131	LPALAGTCHRVMTIVPRYDQYKDAWDSVLEVKVGDRTKRVFFHCYKRGVDRVFDH	190	
QY	140	PWFLAKVWGKTKSLYGPGRGADYLDNKHKRFALFCYKAATEAARVLPF-----GP-GED	191	
Db	191	PWFLEKVGKTKQKLYGPTGTNDYEDNQRLSFQAALEAPRVLUSLSSKSYFGPYGED	250	
QY	192	CVFVANDHSAIVPVLKDEYQPKGQFTKAKSVLAHNIAFOGRMWEBAFKDKLPQAAF	251	

Db 251 VIFVANDWHTALIPCYLKSMTQSRGIYTNARVVFCHINATYQGRFAFADFSLLNLPDOFK 310
QY 252 DKLAESDGYAKYVTEATPMEDEKXPLGKTKYKKNLWKGIIAADKLVTSVPNATETA 311
Db 311 SSFDFIDGHVK-----PVVG---RKNLWLAGLIESFVITVSPNAYAKELV 353
QY 312 ADAAGVELDVTIRA-----KGIEGIVNGMDIEWNPKTDKFLSAPYDQNSVYAGKAAAKE 367
Db 354 SGPDFGVELDNIIRKIDDDGRLVGVNGMDVOEWNPETDKYIAVKYDVSTVLEAKALLKE 413
QY 368 ALQAEGLGVPDPTAPLAFAGFIRLEEQGVDIILALPKLILATPKVQIAILGTGKAAYEKL 427
Db 414 ALQAEVLEGVDRNIPLIGFIRLEEQGVSDILAEALPOFI-KONVOLVALGTGKKOMEQ 472
QY 428 VNAIGTKYKRAKGVVKKFAPLAHMLTAGADFMVPSRPEPGLIOLHAMHYGTVPVVAS 487
Db 473 LEELEISYPDKARGVAKFNVPLAHMLIAGADFLVPSRPEPGLIOLQAMRYGTVPVVAS 532
QY 488 TGLVDTVKEGTVGFHMGALNP--DKLDEADALAAATVRRASEVFAGGRYPPEMVANCIS 545
Db 533 TGLVDTVKEGTVGFHMGALNP--DKLDEADALAAATVRRASEVFAGGRYPPEMVANCIS 545
QY 546 QDLWSKPAQKKEGLEEVVY-----GKGG--VATAKKEIKVP 582
Db 593 QDLWSKPAKAEWEVLLSLGVGSEPGSDGEETAPQAKENVATP 636

RESULT 2
US-10-138-075-2
; Sequence 2, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10/138,075
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Zea mays
US-10-138-075-2

Query Match 39.8%; Score 1449.5; DB 9; Length 609;
Best Local Similarity 48.7%; Pred. No. 2.4e-94;
Matches 302; Conservative 85; Mismatches 164; Indels 69; Gaps 14;
QY 6 TSPSSARPVINAASEGVKKTANOLLRELARGS-----ARKSTSRSAVTGATG- 54
Db 16 TNRESALK---QAPHMOFOCCNGGLRFLSKHSQSRSKIQVAKRRATDNGIHPKTTGH 71
QY 55 -----ATCADDIVWAAEVAPWSKTGGLGDTVGLPIELVKRGRHVTIAPRYDOYADAW 109
Db 72 RAPIVCSAGMTIVEVATEVHPWCKTGGLGDDVVGGLPPALAAAGHRVMTIAPRYDOYKDAW 131
QY 110 DTSVVDI-MG---EKRYFHSIKKGVHRVMDHPFLAKVWGTGSKLYGPRSGADYLD 165
Db 132 DTSVLVEVNGDVTETVEFFCYKRGVDRVDFVHPMFLKVMGKTGAKLYGTTGTDYRD 191
QY 166 NHRFALPCKAAAEARVLPF-----GP-GEDCVFVANDWHSALVPLKDEYQPKGQ 217
Db 192 NOLRCELCLAALEAPRVLNENNSYFSGPYGDFVDFVANDWHTAILPCYLKSMYKPKNGI 251
QY 218 FTKAKSVLAHINIAFGRMWEEAFKDTLPQAAFDKLAFSOGYAKVYVTEATPMEDEKPP 277

Db 252 YNNKAVAFCHINATYQGRFAFADFSLLNLPDSEFLPSDFIDGHVK-----P 297
QY 278 LTGKTKYKKNLWKGIIAADKLVTSVPNATETAADAAGVELDVTIRAKGIB-GIVNGM 336
Db 298 VLG---RKNLWKGIIESDLVTSVPNATETAADAAGVELDVTIRAKGIB-GIVNGM 354
QY 337 DIEWNPKTDKFLSAPYDQNSVYAGKAAKEALQAEGLGVPDPTAPLAFAGFIRLEEQGV 396
Db 355 DVEWNPSTDKYISAKYDATVTEARALNKRERLQAEVGLPVDSSIPVIFVGRLEEQKS 414
QY 397 DIILALPKLILATPKVQIAILGTGKAAYEKLVAIGTKYKRAKGVVKKFAPLAHMLTAG 456
Db 415 DIILALPEFVG-ENVQIIVIGTGKKEEELQLLEVKNPNNARGIAKFNVPPLAHMMFAG 473
QY 457 ADEMLVPSRPEPGLIOLHAMHYGTVPVVASSTGGLVDTVKEGTVGFHMGALNP--DKLDE 514
Db 474 ADLIVPSRPEPGLIOLQAMRYGTVIPICSSSTGGLVDTVKEGTVGFHMGFNFVCECTVP 533
QY 515 ADADALAAATVRRASEVFAGGRYPPEMVANCISQDLWSKPAQKKEGLEEVVYKGG----- 569
Db 534 ADVTAVASTVTRALKQYDTPAFHEMVONCMAQDLWSKGPAAKW-----EVELLGLGVBSR 589
QY 570 -----GVATAKKEIKVP 582
Db 590 AGIDDAEIEIAPLAKENVATP 609

RESULT 3
US-10-138-075-5
; Sequence 5, Application US/10138075
; Publication No. US20030087369A1
; GENERAL INFORMATION:
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: Granule-Bound Starch Synthase
; FILE REFERENCE: BB1474 NA
; CURRENT APPLICATION NUMBER: US/10/138,075
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/288,315
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-138-075-5

Query Match 38.1%; Score 1388.5; DB 9; Length 599;
Best Local Similarity 48.5%; Pred. No. 4.8e-90;
Matches 288; Conservative 84; Mismatches 167; Indels 55; Gaps 13;
QY 18 NAASEGVKKTANOLLRELAR--GSARKSTSRSA-VTGATGATCALDIVWAAEVAPWSKT 74
Db 32 NEVMFLSMRNTKQAKRRATNYGTHRNSSRPAPIVCSG---MPIIFATEVHPWCKT 87
QY 75 GGLGDTVGLPIELVKRGRHVTIAPRYDOYADAWDTSVVVDIM---GEKRVYFHSIKK 130
Db 88 GGLGDTVGLPPALAAAGHRVMTIAPRYDQKDTWDTNVLVEVIVGDRTEVRFPHCYKR 147
QY 131 GVHRVMDHPFLAKVWGTGSKLYGPRSGADYLDNHRKRFALPCKAAAEARVLPF----- 186
Db 148 GYDRVDFVHPMFLKVMGKTGSKLYGTTGTDFRDNQLRFLCLLCLAALEAPRVLNINSE 207
QY 187 ---GP-GEDCVFVANDWHSALVPLKDEYQPKGQFTRAKSVLAHINIAFGRMWEEAFK 242
Db 208 YFSGYGENVVFVANDWHTAVLPCYLKSMYKONGIYVNAKVAFCILHINATYQGRFVDFE 267
QY 243 DTKLPOAAFDKLAFSOGYAKVYVTEATPMEDEKPPDTGKTKYKKNLWKGIIAADKLVTV 302

Db 268 LNLPSFMSFDFVDGHVK-----PVVG-----RKINWAKGITECDVVLIV 310
QY 303 SPNATEIRADAAGVELDVIIRAKGIE-GIVNGMIDIEWNPXTDKFELSAPYDONSYYAG 361
Db 311 SPHYVELTSGPKGVELDGLVRAKPLETGIVNGMDVVDWNPATDKYISVKYNATTVAEA 370
QY 362 KAAAKALQALGLPVDPTAFATFGRLEEQKGVDIILAAALPKILATPKVQIATLGTGK 421
Db 371 RALNKEILOAEVLFPVDSSTPIVIFGRLEEQKGVDIILAAALPEFL-EENVQIIVLGTGK 429
QY 422 AAYEKLVNAGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFPFCGLIOLHAMHYGT 481
Db 430 KMEELMLLEAKYQNGARCIKFNVPPLAHMMFAGANFIIVPSRFPFCGLIOLQGRYGV 489
QY 482 VPVASTGGLVDFVKGSGVGFHMGALNP--DKLDEADALAAATVRRASVFFAGGRYPEN 539
Db 490 IPICSTGGLVDFVSEGVTFHMGSNFVEFETVDPADVAASVNVTRALKQYKTPSFHAM 549
QY 540 VANCISQDLWSKPAOKWEGILLEEVYVYKGG-----GVATAKKEIKVP 582
Db 550 VQNCMAQDLWSKPAKKW---EAILGLGVBSQPGIEETIAPLAKONVATP 599

RESULT 4

US-09-952-677-2
; Sequence 2, Application US/09952677
; Patent No. US20020138876A1
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Lorz, Horst
; Lutticke, Stephanie
; Walter, Lennart
; Froberg, Claus
; Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; FROM WHEAT WHICH ARE INVOLVED IN STARCH
; SYNTHESIS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:
ADDRESS: James F. Haley, Jr., c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/952,677
FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/196,390
FILING DATE: 19-No. US20020138876A1-1998
APPLICATION NUMBER: DE 196 21 588.9
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: DE 196 36 917.7
FILING DATE: 11-SEP-1996
APPLICATION NUMBER: PCT/EP97/02793
FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Haley, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: AGREVO-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 671 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-952-677-2

Query Match 24.3%; Score 884.5; DB 10; Length 671;
Best Local Similarity 34.3%; Pred. No. 2.3e-54;
Matches 248; Conservative 98; Mismatches 245; Indels 133; Gaps 24;

QY 23 GVKTAQLLRELARGSARKSTSRSAVTGATGATCALD-----IVMVAEAVP 70
Db 9 GVSELAPDLLE---GIAEDSIDSIIVAASEQDSEIMDANEQPOAKYTRISFVFTGAA 65
QY 71 WSTGGLGDVGTGPIELVKRGHVMITAPRY-----DOYADAWDTSVVVDIM---GE 120
Db 66 YAKSGGLGDVCGSLPITALAARGHVMVMPRYLNGSSDKNYAKALYAKHKKIPCTCGSH 125
QY 121 KVRYPHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKFALEKAAIEA 180
Db 126 EVTFEYHRDNDVDFVDPH-----SYHRPGS-LYGDNFGA-FGDNQFRYTLTLLCYAA 178
QY 181 ARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKQFTKAKSVLAHNTAFQGRMW 237
Db 179 PLILELGGYIYGQNCMFVNDWHSALVPVLLAARYGYVYRDSRSTLVTHNLAHQVGP 238
QY 238 EEAFKDTKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTGTGYKKNLWKGIIAAD 297
Db 239 ASTYPDGLPPEWVGALEW-----VPEWARHRLDXG-----EAVNFLAGAVVTAD 285
QY 298 KLVTVSPNATEITAADAAGVVELDVI--RAKIGIEIVNGMDTEENPKTDKFLSAPYDQ 355
Db 286 RIVTVSQYSGYSEVTT-AEGQGLNELLSRKSRLNGIVNGIDINDWNPITDKCLPHYSV 344
QY 356 NSVYAKAAKAEALQALGLPVDPTAFATFGRLEEQKGVDIILAAALPKILATPKVQIA 415
Db 345 DDL-SGAKCKAELOKELGLPVREDVPLIGTIGRDIYQKIDILIKWAIPELM-REDVQV 402
QY 416 ILGTGKAAVEKLVNAGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFPFCGLIQLH 475
Db 403 MLGSGDPITFEGWNRKSTESSYKDKFRGWGVSVPSVSHRITAGCDILLMPSEFECGLNQLY 462
QY 476 AMHYGTVPVASTGGLVDIVK-----EGVTFHMGALNPKDLDEADALAAATVRR 526
Db 463 AMOYGTVPVYVHTGGLRDIETFPNPFCAKGEETGWAFLPTVYDKMLWA-----LRT 514
QY 527 ASEVFAAGGRYPENVANCISQDLWSKPAOKWEGILLEEVYVYKGVATATKEELKVPVAK 586
Db 515 AMSTFE-----HAPS--WEGLM-----KRGWTKDHTWDH 542
QY 587 IPGDLPAVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSAAAT---P 643
Db 543 APSSTSRSSSGPSWTNPTSC-----RRGLRSKCESPSALKTSSSSFRGP 587
QY 644 KVTYKPAALPATAKPTAGLK-LAGEASTTSENGAASNGNGNGNGASAKTSAKPLVSA 702
Db 588 EGYPTLCRPAIVESOCACLLWFAG-----SRTYDGA-----AAAVTASGGROLPW 635
QY 703 ATRK 706
Db 636 GIRK 639

RESULT 5

US-09-952-677-6
; Sequence 6, Application US/09952677
; Patent No. US20020138876A1
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Lorz, Horst
; Lutticke, Stephanie
; Walter, Lennart
; Froberg, Claus
; Kossmann, Jens

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
FROM WHEAT WHICH ARE INVOLVED IN STARCH
SYNTHESIS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/952,677

FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/196,390

FILING DATE: 19-No. US20020138876A1-1998

APPLICATION NUMBER: DE 196 21 588.9

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: DE 196 36 917.7

FILING DATE: 11-SEP-1996

APPLICATION NUMBER: PCT/EP97/02793

FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Haley, Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: AGREVO-9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 799 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-952-677-6

Db 624 GKGVEIITADAMPWIV-SQDVQLVMIGTGRHDESMLOHFREHDDKVRGWGFSVRLAH 682
QY 452 MLTAGADEMLVPSREFPCGLIOLHAMHYGTVPVWASTGGLVDTV-----KEGVTGFMHG 505
Db 683 RITAGADALLMPSREFPCGLNOLYAMAYGTVPVWAVAGLRTVTPFPDFNFHSGLGWTFD 742
QY 506 ALNPKLDEADADALAAATVRRASEVAGGRYPFEMVANCISQDLSWSKPAQKWEGLLEVV 565
Db 743 RAFAHKLIEALGHCL-RTYRDFKE-----SWRALQERGMQSQDSNEHAAKLYEDVLVRAK 796
QY 566 Y 566
Db 797 Y 797

RESULT 6

US-10-163-214-6

; Sequence 6, Application: US/10163214

; Publication NO. US20030097688A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen M.

; APPLICANT: Broglie, Karen E.

; APPLICANT: Butler, Karlene H.

; APPLICANT: Thorpe, Catharine J.

; TITLE OF INVENTION: Starch Synthase Isoform V

; FILE REFERENCE: BB1520 US NA

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US/10/163,214

; PRIOR FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6

; LENGTH: 915

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-163-214-6

Query Match 15.4%; Score 562.5; DB 9; Length 915;
Best Local Similarity 29.0%; Pred. No. 2a-31;
Matches 173; Conservative 94; Mismatches 199; Indels 131; Gaps 22;

QY 32 LRELARG-----SARKSTSRSAVTC-----ATGATCALTDIVVAEAVPWSKGG 76
Db 380 LREMVRKRDSCLEAYILSCRMKMDREIVDNFLKITLPCTSSGLHIHIAEAMAPVAKVG 439
QY 77 LGDVTGSLPIELVKRGHRYMTIAPRYD-QYADAWTSVVVDIMGEKVRYPHSIKKGVRHV 135
Db 440 LADVISGLKALOKKHLVEIILPKYDCMONQVNNLKVLDDVVVQS--YF-----487
QY 136 WIDHPWFLAKVWGKT--GSKLY--GPRSGADYL-----DNHKKRFALFCAAIEAAR 182
Db 488 --EGNLFNNKNIWTGTVGSLPVYFTEPQHPAKFFWRAQYGEHDDFKRFAYFSAALE---542
QY 183 VLPFGGECDCVFNADWHSALVPVLLKDEYQPKQFTKAKSVLAHIAHIAQGRWBEAFK 242
Db 543 LLYQSKKIDIIHCHDMQTAFAVAPLYWEAYANLG-FNSARICTCHNFYQG-----593
QY 243 DTKLPGAADFALFSDGYAKVYTEATPMEE-----DEKPPITGTYTKKINWLGK 291
Db 594 -----AAPADLACGLDVOQLDREDMRDNHSHGRINNVKG 629
QY 292 GIITADKLVTSPNYATEIAADAGGVVELDTVIRAKGIEGVINGMDTEENPKTDKFLSA 351
Db 630 AIVSVNITVTSPYIALEVRSEGRGLQDSLKHSRKFGVGLINGIDTDTWNPSTDRHLK 689
QY 352 PYDONSVAKAAKAEALQAEGL-PVDPTAPLFAFTGRLEEOKGVDIILALPKIATP 410
Db 690 QYANDL-QGKAANKAARLNKLNLSSTNASOPLVGCITRLVPQKGVHLIRHAIYK-TAEL 747
QY 411 KVOITAILGTCKAAEYKLVNAIGTKYKGRA-----KGWVKFSAPLAHMLTAGADFML 461

QY 50 TGATGATCALTDIVVAEAVPWSKGGTGLGDTGGGLPIELVKRGHRYMTIAPRYDQYADAW 109
Db 299 SGPLAGENVNVVVAEECPWCKTGGTGLGDTGALPKALAKRGHRYMVVVPYGYEEAY 358
QY 110 DTSV--VVDIMGE--KVRYPHSIKKGVRHVWIDHPWFLAK---VWCKTSGSKLYGPRSGAD 162
Db 359 DVGVRKYKAAQODMEVNYFAYIDGDVDFVFIADPLFRHQEDYIGGSRQEI-----411
QY 163 YLDNHRKRFALFCAAIEA-----ARVLPFGGDCVFNADWHSALVPVLLKDEYQPKQ 217
Db 412 -----KRMILFCKAAVEVPHVPCGVPYGDG-NLVFIANDWHTALLPVLYKAYYRDHGL 465
QY 218 FYTAKSVLAHNTAFQGRWBEAFKDTKLPGAADFALFSDGYAKVYTEATPMEEDEKPP 277
Db 466 MQYTRSIMVLIHIAHQGRGVPDFEFTPELPEHLEHFLYD-----P 507
QY 278 LICKYTKKNLWLGKIIADKLVTSPNYATEIAADAGGVVELDTVIRAKG--IEGVING 335
Db 508 VGG---EHANYFAAGLKMDQVVVSPGYLWELKT-VEGGWGLHDIIRONDWKTGRVING 563
QY 336 MDTEENPKTDKFLSAP-----YDONSVAKAAKAEALQAEGLGPDPTAPLFAFTGRLE 391
Db 564 IDNWNENPEVDALHKSQYTNFSLRTLDGSKQCKEALQRELGLQVRADVPVLLGFIGRLD 623
QY 392 EQKGVDIILALPKIILATPKVQITAILGTCKAAEYKLVNAIGTKYKGRAKGVVKSAPLAH 451

Query Match 23.1%; Score 842; DB 10; Length 799;
Best Local Similarity 38.6%; Pred. No. 3a-51;
Matches 209; Conservative 76; Mismatches 190; Indels 66; Gaps 15;

```

Db 748 GGQFVLGSSFPVPH-----IQREFEGIAHFQNNNIRLLLLKYDDSLSHWIAASDMFI 801
QY 462 VPSRFEPCGLIQLHAMHYGVPVVASGGLVDTV-----KEGVTGFHMGALNPKL 512
Db 802 VPSWFPCGLTQMAMRYGSPVVRKVTGGLNDSYFDFDDETPKELRNGFTF--VHPDE- 858
QY 513 DEADALAAATVRRASVFAGGRYPENVANCISODL-----SWSKPAOKWEGILLEEV 565
Db 859 -----KALSGAMERAFNY--NRKPEVWKOLVQKMDRIDFSWASSASQYEDITQRAV 908

RESULT 7
US-10-163-214-2
; Sequence 2, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 909
; TYPE: PRT
; ORGANISM: Zea mays
US-10-163-214-2

Query Match 15.4%; Score 559; DB 9; Length 909;
Best Local Similarity 28.8%; Pred. No. 3.6e-31;
Matches 174; Conservative 98; Mismatches 218; Indels 114; Gaps 23;

QY 10 SSARPVINAASFVGVKKTANOLLRELA-----RGSARKSTSRSAVTGA-TG 54
Db 361 SLEKKIINDAS-----MLREMAWKEDNKLREAYLSRGMEERELIDSPKMLPG 411
QY 55 ATCGLDIVMVAEEVAPWSKGTGGVDTGGLPTLKVKGHRVMTIAPRYD-QYADAWDTSV 113
Db 412 TSSGLHIVHIAEMAPVAKVGGGLADVISGLGKALQKGLHVEILPKYDCMOHQNINLK 471
QY 114 VVDIMGEKRYPHSTIKKGVRHWIDHPWFLAKWGKT--GSKLY--GPRSGADYL----- 164
Db 472 VLDV-----VVKSYFEGNMFANKIWTGTVGGLPVYFTEPQHPGKFFWRAQY 517
QY 165 ----DNHKKRALFCAAATEAARVLPFGGEDC-VFVANDWHSALVPVLLKDEYQPKGQFT 219
Db 518 YGEHDDKRFYSFVSVALE-----LTYQSGKKVDIIHCHDWQAFVAPLYWDVYANLG-FN 572
QY 220 KAKSVLATHNIAFOGRMWEAEAFKTLPOAFAFKLAFSDGYAKVYVTEATPMEDEKPLT 279
Db 573 SARICFTCHNFYQG-----IAPAQDLAYCGLDVDHLDORPDR--MR 611
QY 280 GKTYKKNLWKGIIIAADKLVTVPSPNYATEIAADAGGVVELDTVIRAKGIEGVINGMOIE 339
Db 612 DNSHGRINVGKAVVYSNIVTVPSTYAOEVSEGGRLQDITLVKHSKFFVGILNGIDTD 671
QY 340 ENNPXTDKFLSAPYQNSYAGKAAKALQALGL-PVDPTAPLFAFAGRLEEGKGVDI 398
Db 672 TNPSTDFRLKVOYSANDLY-GKSANKALRQLKLASTQASQPLVGCITRLVPOKGVHL 730
QY 399 ILAALPKILATPKVQIAILGTGKAAK-EKLVNAGITKYK--GRAGGVVKSAPLAHMLTA 455
Db 731 INHAIYKIKTELGG-QVLLGSSPVQHIQREFEGIAHQFQNNNNVRLLLKYDALAHMIFA 789
QY 456 GADEMLVPSRPEPCGLIQLHAMHYGTVPVVASTGGLVDT-----VKEGVTGFH 503
Db 790 ASDMFI VPSMEPCGLTQWAMRYGSPVVRVTGGLNDSYFDFDDETPIMPEVRNGFTFL- 848

```

```

QY 504 MGALNPDKLDRADADALAAATVRRASVFAGGRYPENVANCISODL-----SWSKPAOKWEG 559
Db 849 -----KADBDQ-----FGNALERAFNY--HRKPEVWKQLVQKMDRIDFSWDTSVSQYEE 896
QY 560 LLEE 563
Db 897 IYQK 900

```

RESULT 8

```

US-10-163-214-12
; Sequence 12, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-163-214-12

Query Match 15.3%; Score 555.5; DB 9; Length 914;
Best Local Similarity 29.8%; Pred. No. 6.4e-31;
Matches 176; Conservative 96; Mismatches 216; Indels 103; Gaps 24;

```

```

QY 31 LRELA-----RGSARKSTSRSAVTGA-TGATCALDIVMVAEEVAPWSKGT 75
Db 378 MLREMYQKDRSLREAYLSYRGTEENREVMDNLLKMLALPGTSSGLHTAHIAEMAPYAKVG 437
QY 76 GLGDTVGGPLPIELVKKGRHVMTIAPRYD-QYADAWDTSVVDIMGEKRYFHSIKKGVHR 134
Db 438 GLADVTSGLGKALQKGLHVEILPKYDCMOVDSNLKVLVDLVQS--YF----- 486
QY 135 VWIDHPWFLAKWGKT--GSKLY--GPR-----SGADYLDNH--KRALFCAAALEA 181
Db 487 ---EGNMFNKNITGIVEGLPVYFIEPQHPAMFFSRAQYGEHDDKRFYSFRAALE-- 541
QY 182 RVLPPFGGEDC-VFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLATHNIAFOGRMWEA 240
Db 542 --LLYQSGKKVDIIHCHDWQAFVAPLYWDVYANLG-FNSARICFTCHNFYQG----- 592
QY 241 FKDTKLPOAFAFKLAFSDGYAKVYVTEATPMEDEKPP-LTGKTYKKNLWKGIIIAADKL 299
Db 593 -----TAPARDLAWC-----GLDVEHLDRPDRMRDNRSHGRINAVKGAVVYSNIV 636
QY 300 VTVPSPNYATEIAADAGGVVELDTVIRAKGIEGVINGMOIEENPKTDKFLSAPYQNSYV 359
Db 637 TTVSPTYALEVRSEGGRLQDITLVKHSKFFVGILNGIDTDWNPCPTRYLKVOYNKADL- 695
QY 360 AKSAAKAEALQALGL-PVDPTAPLFAFAGRLEEGKGVDIILAAALPKILATPKVQIAILG 418
Db 696 QGKAANKAALRQLKLASTQASQPLVGCITRLVQKGVHLIRHAIYK-TAELGGQFVLIG 754
QY 419 TCKAAEYKLVNAGITKYKGRA-----KGWVFSAPLAHMLTAGADFMVPSRFEPC 469
Db 755 SSP-----VPEIQREFEGIAHFQNNNIRLLKYDDALSCHIYAAASDMFIVPSIFEPC 808
QY 470 GLIQLHAMHYGTVPVVASTGGLVDTV-----KEGVTGFHMGALNPKLDRADADALAAATVR 525
Db 809 GLTQMAMRYGSPVVRVTGGLNDSYFDFDDETPIMPEVRNGFTFVKADESGQLSSAME 865

```



```
; SEQ ID NO 69
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-69

Query Match      11.1%; Score 402.5; DB 12; Length 476;
Best Local Similarity 26.9%; Pred. No. 1.7e-20;
Matches 150; Conservative 78; Mismatches 204; Indels 125; Gaps 24;

QY 61 IVWVAEAPMSKGGGLDVTGGPIELVKGHRVMTIAPRY---DOVADAWDTSVVVDI 117
Db 3 IVQVAVEFTPIVKVGGGLDVAASLSKELAKQ-NDVEVLLPHYPLISFESS-----QV 54
QY 118 MGKVRVFHSKKGVHRVWDHPWFLAKVWCKTGSKLYGP-----RSGAD 162
Db 55 LSENSFYE-----FLGKQQAASALSISYEGTLTIIITLDSQIELFSTTSV 99
QY 163 YLDNH-KRFALFCKAAIEAARVLFPFGEDCVFVANDWHSALVPVLKDKDEQPKGQTKA 221
Db 100 YSENNVVFSAFAAAA--AAYLQADPAD--IVHLHDWHVGLLAGLLKNPLNP-----VHS 151
QY 222 KSVLAHINIAFQGRWEEAEKDTKLPOAFDKLAFSDGYAKVYTEATPMEDEKPLTGK 281
Db 152 KIVFTIHNFGYRG-----YCSSTOLLAAS-----QIDDFHLSHY-----QLFRDPQTSV 194
QY 282 TYKKNLWKGIIIAADKLVTSPNRYATEIAADAAGGVYELDTVIRAKG--IEGIVNGMDIE 339
Db 195 L-----MKGALYCSDIITVSTYVQEIINDYS-DYELHDAILARNSVFSGIINGIDED 247
QY 340 ENPKTKDFLSAPYD-----QNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQ 393
Db 248 VWNPKTDPALAVQYDASLLSEPDVLTKEENRAVLVEKLGISD-YFPLICVLSRIVEE 306
QY 394 KG-----VDIILALPKLATPKVOIALGTGKAAEKLVAIGTKYKGRAG-----VVK 444
Db 307 KGPEFKKEIILHMEHSYA-----FILGTSQ--NEVLNFEFRMLQDCLASSPNRILLD 359
QY 445 FSAPLAHMLTAGADEMLVPSRFPCCGLIQLHAMHYGTVPVASTGGGLVDTVKGVGFHM 504
Db 360 FNDPLARLYAADMICIFSHREACGLTQLIAMRYGTVPVLRKGTGGGLADTVIPGVNGFTF 419
QY 505 GALNPDKLDADALAAATVRRASEVFAAGRYPEMVANCISODLSKSPAKKQEGGLEEV 564
Db 420 --FDTNNFNE-----FRMLSNAYT---TYRQEPDWMNLIESG 453
QY 565 VYKGGVATAKKEIKV 581
Db 454 MLRASGLDAMAKHYVNL 470

RESULT 12
US-10-163-214-10
; Sequence 10, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-163-214-10

Query Match      11.1%; Score 402.5; DB 12; Length 476;
Best Local Similarity 26.9%; Pred. No. 1.7e-20;
Matches 150; Conservative 78; Mismatches 204; Indels 125; Gaps 24;

QY 61 IVWVAEAPMSKGGGLDVTGGPIELVKGHRVMTIAPRY---DOVADAWDTSVVVDI 117
Db 3 IVQVAVEFTPIVKVGGGLDVAASLSKELAKQ-NDVEVLLPHYPLISFESS-----QV 54
QY 118 MGKVRVFHSKKGVHRVWDHPWFLAKVWCKTGSKLYGP-----RSGAD 162
Db 55 LSENSFYE-----FLGKQQAASALSISYEGTLTIIITLDSQIELFSTTSV 99
QY 163 YLDNH-KRFALFCKAAIEAARVLFPFGEDCVFVANDWHSALVPVLKDKDEQPKGQTKA 221
Db 100 YSENNVVFSAFAAAA--AAYLQADPAD--IVHLHDWHVGLLAGLLKNPLNP-----VHS 151
QY 222 KSVLAHINIAFQGRWEEAEKDTKLPOAFDKLAFSDGYAKVYTEATPMEDEKPLTGK 281
Db 152 KIVFTIHNFGYRG-----YCSSTOLLAAS-----QIDDFHLSHY-----QLFRDPQTSV 194
QY 282 TYKKNLWKGIIIAADKLVTSPNRYATEIAADAAGGVYELDTVIRAKG--IEGIVNGMDIE 339
Db 195 L-----MKGALYCSDIITVSTYVQEIINDYS-DYELHDAILARNSVFSGIINGIDED 247
QY 340 ENPKTKDFLSAPYD-----QNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQ 393
Db 248 VWNPKTDPALAVQYDASLLSEPDVLTKEENRAVLVEKLGISD-YFPLICVLSRIVEE 306
QY 394 KG-----VDIILALPKLATPKVOIALGTGKAAEKLVAIGTKYKGRAG-----VVK 444
Db 307 KGPEFKKEIILHMEHSYA-----FILGTSQ--NEVLNFEFRMLQDCLASSPNRILLD 359
QY 445 FSAPLAHMLTAGADEMLVPSRFPCCGLIQLHAMHYGTVPVASTGGGLVDTVKGVGFHM 504
Db 360 FNDPLARLYAADMICIFSHREACGLTQLIAMRYGTVPVLRKGTGGGLADTVIPGVNGFTF 419
QY 505 GALNPDKLDADALAAATVRRASEVFAAGRYPEMVANCISODLSKSPAKKQEGGLEEV 564
Db 420 --FDTNNFNE-----FRMLSNAYT---TYRQEPDWMNLIESG 453
QY 565 VYKGGVATAKKEIKV 581
Db 454 MLRASGLDAMAKHYVNL 470

Query Match      10.6%; Score 385.5; DB 9; Length 293;
Best Local Similarity 33.0%; Pred. No. 1.4e-19;
Matches 101; Conservative 59; Mismatches 115; Indels 31; Gaps 9;

QY 285 KINLWKGIIIAADKLVTSPNRYATEIAADAAGGVYELDTVIRAKGIEGIVNGMDIEWNP 344
Db 1 RINAVKGAVYSNIVTTSPTVYALVEVRSEGGRLQDTLKVHSRKFGLIINGIDITWNP 60
QY 345 TDKLSADYDQNSYVAGKAAAEALQAEGL-PVDPTAPLFAFGRLEEQKGVVDIILAA 403
Db 61 TDRVLKVOYNAKDL-QGKAANKAALREQLNLSAYSPQPLVGCITRLVAQKGVHLIRHAI 119
QY 404 PKILATPKVQIAILGTGKAAEKLVAIGTKYKGRAG-----KVVKFSAPLAHMLT 454
Db 120 YK-TAELGGQFVLGGSP-----VPEIQREFEGTADHFQNNNIRLLIKYDDALSHCIY 172
QY 455 AGADFMVPSRFPCCGLIQLHAMHYGTVPVASTGGGLVDTV---KEGVTGFHMGALNP 510
Db 173 AASDMFIVPSIFEPCCGLTQMIAMRYGSPVIVRKTGGGLNDSVDFDEDETIP---MEVRNGF 229
QY 511 KLDEADADALAAATVRRASEVFAAGRYPEMVANCISODL---SWSKPAKQKWEGLLEEV 566
Db 230 TFVKADEGGUSSAMERAFNCYT--RKPEVWKOLYKQDMTIDFSWDTASQYEDIYQKAVA 287
QY 567 KGKGA 572
Db 288 RARAVA 293

RESULT 13
US-09-739-438-2
; Sequence 2, Application US/09739438
; Patent No. US20020029394A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Beckles, Diane M.
; APPLICANT: Thorpe, Cathy
; TITLE OF INVENTION: Homologs of Starch Synthase DUI
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/739,438
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/171514
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-739-438-2

Query Match      5.9%; Score 215.5; DB 10; Length 117;
Best Local Similarity 46.6%; Pred. No. 3.9e-08;
Matches 41; Conservative 17; Mismatches 25; Indels 5; Gaps 1;

QY 413 QIALGTG-----KAAVEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADEMLVPSR 467
Db 8 QVVLGSAPDHRIQDFTNLASKLHGEYHGRVKLCITYDEPLSHLIYAGADFILVPSMFE 67
QY 468 PCGLIQLHAMHYGTVPVASTGGGLVDTV 495
Db 68 PCGLTQLTAMRYGSPVIVRKTGGGLVDTV 95

RESULT 14
US-09-738-626-4738
; Sequence 4738, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
```


GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:49:45 ; Search time 34.5277 Seconds
(without alignments)
2512.365 Million cell updates/sec

Title: US-09-980-771a-3_COPY_58_708

Perfect score: 3370

Sequence: 1 ALDIVMAAEVAPWSKTGL.....SASKTSAKPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	42.7	606	13 AAR25476	Granule-bound star
2	1397.5	41.5	509	19 AAW56485	Oryza sativa starch
3	1390	41.2	527	21 AAG04668	Arabidopsis thalia
4	1390	41.2	563	21 AAG04667	Arabidopsis thalia
5	1390	41.2	610	21 AAG04666	Arabidopsis thalia
6	1390	41.2	610	23 ABB91229	Arabidopsis thalia
7	1353.5	40.2	533	19 AAW56484	Herbicidally activ
8	1344.5	39.9	637	19 AAY16604	Zea mays waxy gene
9	1331.5	39.5	502	13 AAR25474	Protein encoded by
10	1194	35.4	238	22 AAB31175	Central fragment f
					A granule bound st

11	904	26.8	534	21 AAB49307	Wheat starch synth
12	898	26.6	792	23 ABB92160	Herbicidally activ
13	891.5	26.5	647	20 AAY09004	Wheat starch solub
14	882.5	26.2	649	19 AAW38218	Maize starch synth
15	881	26.1	671	19 AAW23937	Wheat starch synth
16	881	26.1	756	21 AAY50818	Wheat soluble star
17	873.5	25.9	626	15 AAR51231	Wheat soluble star
18	870.5	25.8	652	23 ABB93595	Soluble rice starch
19	861.5	25.6	539	19 AAW56491	Herbicidally activ
20	861.5	25.6	583	19 AAW56488	Zea mays starch sy
21	852.5	25.3	583	19 AAW70894	Zea mays starch s
22	847	25.1	802	23 AAU99845	Maize starch solub
23	847	25.1	813	23 AAU99844	Modified barley st
24	846	25.1	812	23 AAU99847	Barley cultivar Mo
25	845	25.1	804	19 AAW70892	Barley line MK6827
26	844	25.0	798	21 AAB37566	Maize starch solub
27	842	25.0	799	21 AAB37567	Wheat starch synth
28	842	25.0	799	21 AAB37597	Wheat starch synth
29	841	25.0	799	19 AAW23938	Wheat starch synth
30	837	24.8	698	19 AAW56487	Wheat starch synth
31	825.5	24.5	812	23 AAU99846	Zea mays soluble s
32	825	24.5	597	21 AAB37568	Barley line 292 st
33	767	22.8	466	23 AAM51855	Wheat starch synth
34	726	21.5	459	17 AAR99540	Rice starch synth
35	701.5	20.8	669	19 AAW70893	Soluble starch syn
36	701.5	20.8	669	19 AAW56486	Maize starch solub
37	669	19.9	483	23 ABB54012	Zea mays soluble s
38	647.5	19.2	476	23 ABB27592	Lactococcus lactis
39	580	17.2	477	13 AAR25462	Streptococcus poly
40	580	17.2	477	19 AAW53890	giga. Escherichia
41	568.5	16.9	677	19 AAR99539	E. coli glycocon s
42	568.5	16.9	1230	18 AAW17785	Soluble starch syn
43	568.5	16.9	1230	21 AAB49306	Potato tuber solub
44	561.5	16.7	1025	23 ABB90967	Potato starch synt
45	546	16.2	495	19 AAW70885	Herbicidally activ
					Protein encoded by

ALIGNMENTS

RESULT 1
AAR25476
ID AAR25476 standard; Protein; 606 AA.
XX
AC AAR25476;
XX
DT 15-JAN-1993 (first entry)
XX
DE Granule-bound starch synthase of potato.
XX
KW GBSS gene; amylopectin; amylose production;
KW inhibition; promoter; antisense construct.
XX
OS Solanum tuberosum.
XX
PN WO92111376-A.
XX
PD 09-JUL-1992.
XX
PF 20-DEC-1991; 91WO-SE00892.
XX
PR 21-DEC-1990; 90SE-0004096.
XX
PA (AMYL-) AMYLOGENE HB.
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
XX
DR WPI; 1992-250096/30.
XX
DR N-PSDB; AAR26404.
XX
PT Modifying potatoes to form amylopectin starch - using an
PT anti-sense construct to inhibit granule-bound starch synthase

Db 545 KRAIKVVGTPAYEMVRNCMNQDLSWKGPAAKNWNL--LGLGVAGSAPGEGDEIAPLA 602
QY 519 KEIKVP 525
Db 603 KENVAAP 609

RESULT 3
AAG04668
ID AAG04668 standard; Protein; 527 AA.
XX AAG04668;
AC
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 782.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132566.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 41.2%; Score 1390; DB 21; Length 563;
Best Local Similarity 52.7%; Pred. No. 7.3e-101;
Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;
QY 2 LDIVMAEAPWSTGGTGLGVDVTGGLPIELVKRGRHVMITAPRYDQYADAWDTFVVVDI- 60
Db 37 MSVIFGAEGVPSKTTGGLGVLGPPALAAARGHRVMTICPRYDQYKDAWDTFVVQIK 96
QY 61 MEK---VRYFHSIKGVHVRWIDHPFLAKVWGKTSGLYGRSGADYLDNKRKALFC 117
Db 97 VGDKNVRFPHCYKRGVDRVFDHPIFLAKVVGKTSGLYGRSGADYLDNKRKALFC 156
QY 118 KAAIEAARVLPF-----GP-CEDCVFFVANDHWSALVPVLLKDEYQPKGOFTKAKSVLA 169
Db 157 QAALEAPQVNLNNSKXFSYSGPYGEDVFFVANDHWTALLPCYLKSMQSRGVYNNKVVFC 216
QY 170 IHNIAPQGRWEEAFKDTKLPL---QAADFKLAFSDGYAKVYVTEATPNEDEKPLTGKTY 226
Db 217 IHNIAYQGRFAFDYSILLNLPISFKSFD---FMDGYEK-----PVKG--- 256
QY 227 KKNLWKGGLIADKLVTPSPNYATETADAAGGVELDTVIRAKGLEGVNGMDIEFWNP 286
Db 257 RKNWKAALLEAHRVLTSPYIAQELISGVDGRVELHLYRMTKVTSGIINGMDVQWNP 316
QY 287 KTDKLSAPYDQNSYAGKAAKEALQAEGLFVDPPTAPLFAFIFGRLEEGKGVDTIILAAL 346
Db 317 STDKYIDIKYDITVTDAKPLKEALQAAVGLPVDSDVPVIGFIFGRLEEGKGVDTIILVEAI 376
QY 347 PKILATPKVQIALIGTKKAYEKLVAIGTKYGRAGKGVVYKFSAPLAHMLTAGADFMVLP 406
Db 377 SKFMGL-NVQMVILGTGKKMEQAQILELEKFFGKAVGVAKFNVPPLAHMITAGADFTIIVP 435
QY 407 SRPECGLIQHAMHYGTVPVASTGGLVDTVKREGVTPHMGALNP--DKLDEADADALA 464
Db 436 SRPECGLIQHAMRYGTVPVASTGGLVDTVKDGTYGTFHIGRFNVKCEVVDVDDVATA 495
QY 465 ATVRRASEVAGGRYEMVANCITSQILNSKPAQKWEGLLEEVYCKGVYATAKKEIKY 524
Db 496 KAVTRAVAVYGTSAQEMVKNQMDQFSWKGPARLWEKVLVLLSLNVA-GSEAGTGEETI-A 553
QY 525 PVAEK 529
Db 554 PLAKE 558

RESULT 5
AAG04666
ID AAG04666 standard; Protein; 610 AA.
XX
AC AAG04666;
XX
DT 17-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 780.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 780.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

QY 287 KTDKFLSAPYQNSYAGKAAAKALQALGELPVDPTAPLFAFAGRLGEEQKGVDIILAAAL 346
 Db 364 STDKYIDIKYDITVTDAKPLIKALQAAVGLPVDVDPVVICFGLGEEQKGSILVEAI 423
 QY 347 PKLATPKVQIATLGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVP 406
 Db 424 SKFMGL-NVQMVLGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVP 482
 QY 407 SRFEPCGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 464
 Db 483 SRFEPCGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 542
 QY 465 ATVRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEIKV 524
 Db 543 KAVTRAVAVYGTSAQEMVKNKCMQDQFSWKGPARLWEKVLNVA-GSEAGTEGEEI-A 600
 QY 525 PVAEK 529
 Db 601 PLAKE 605

RESULT 6
 ABB91229 standard; Protein; 610 AA.
 AC ABB91229;
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 440.
 XX Herbicidal; plant; agriculture; herbicide.
 KW Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX WO200210210-A2.
 PN 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX 28-AUG-2001; 2001WO-EP09892.
 PR (FARB) BAYER AG.
 PA Tietjen K, Weidler M;
 PI WPI; 2002-269010/31.
 DR Identifying plant target proteins for herbicidally active compounds,
 XX comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX Claim 5; SEQ ID NO 440; 26lpp + Sequence Listing; English.

XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

XX SQ Sequence 610 AA;

Query Match 41.2%; Score 1390; DB 23; Length 610;
 Best Local Similarity 52.7%; Pred. No. 8.2e-101;
 Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;
 QY 2 LDIVMAAEVAPWSKGTGGLGDTGGLPIELVKGHRVMTIAPRYDQYADAWDTSVVVDI-60

Db 84 MSVIFGAEGVPSKGTGGLGDLVGLPPALAAARGHRVMTICPRYDQYKDAWDTCVVQTK 143
 QY 61 MGEK---VVEFHSIKKGVRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117
 Db 144 VDGKVENVAFHCYKRGVDRFVDHPIFLAKVWGTGSKLYGPRSGADYLDNHRKRALFC 203
 QY 118 KAAIEAARYLPF-----GP-GECDVFVANDHWSALVPLVLLKDEYQPKQFTKAKSVLA 169
 Db 204 QAAIEAARYLPF-----GP-GECDVFVANDHWSALVPLVLLKDEYQPKQFTKAKSVLA 263
 QY 170 IINIAPOGRMEEAFKDTKLP---QAADFKLAFSDGAKYVTEATPEEDEKPLTKTY 226
 Db 264 IINIAPOGRMEEAFKDTKLP---QAADFKLAFSDGAKYVTEATPEEDEKPLTKTY 303
 QY 227 KINLWLGKGIITADKLVTVPNATYATIAADAAGGVDELTVIRAKGTGIVNGMDIEENP 286
 Db 304 RKINWKAALIEAHRVLTVPYQAELISGVDRGVELHKYLRMKTYSGLIINGMDVQENP 363
 QY 287 KTDKFLSAPYQNSYAGKAAAKALQALGELPVDPTAPLFAFAGRLGEEQKGVDIILAAAL 346
 Db 364 STDKYIDIKYDITVTDAKPLIKALQAAVGLPVDVDPVVICFGLGEEQKGSILVEAI 423
 QY 347 PKLATPKVQIATLGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVP 406
 Db 424 SKFMGL-NVQMVLGTGKAAEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADPMLVP 482
 QY 407 SRFEPCGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 464
 Db 483 SRFEPCGLIQLHAMHYGTVPVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALA 542
 QY 465 ATVRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEIKV 524
 Db 543 KAVTRAVAVYGTSAQEMVKNKCMQDQFSWKGPARLWEKVLNVA-GSEAGTEGEEI-A 600
 QY 525 PVAEK 529
 Db 601 PLAKE 605

RESULT 7

AAW56484
 ID AAW56484 standard; Protein; 533 AA.
 AC AAW56484;
 DT 11-SEP-1998 (first entry)
 DE Zea mays waxy gene glucosyl transferase.
 KW SER; starch-encapsulating region; fusion vector.
 XX Zea mays.
 OS Zea mays.
 PN WO9814601-A1.
 XX 09-APR-1998.
 PD 30-SEP-1997; 97WO-US17555.
 PF 30-SEP-1996; 96US-0026855.
 PR (EXSE-) EXSEED GENETICS LLC.
 PA Guan H, Keeling P;
 XX WPI; 1998-240100/21.
 DR N-PSDB; AAV29752.
 XX Hybrid polypeptide comprising starch-encapsulating region and
 PT protein - useful for, e.g. producing protein(s) resistant to
 PT degradation by stomach acids
 XX

Example 2; Page 28; 156pp; English.

The sequence is that of maize glucosyl transferase which is encoded by the waxy gene. It can be used in the production of a hybrid polypeptide comprising a starch-encapsulating region (SER) fused to a payload protein. The hybrid polypeptide can be used to make modified starches comprising the payload protein, selected from, e.g. hormones, growth factors, antibodies, enzymes, dyes, immunoglobulins, etc. The modified starch can also be used to provide grain feeds enriched in amino acids. By encapsulating the payload protein in starch, it is more resistant to degradation by stomach acids.

Sequence 533 AA;

[illegible]

RESULT 8
AAY16604

AB110004
ID AAY16604 standard; Protein; 637 AA.

AA
AC

XX
DT 23-AUG-79

XX
DE

XX
XX
XX

KW starch synthesis enzyme; glycogen synthesis; non-starch branching gene; amylopectin; amylose; plant-like starch.

```

Db 480 FEPCGLIQLOGMRYGTFCACASTGGLVDITTEGKTGFHMGRLSDCNVVEPADYKKVATT 539
AAR25474
QY 467 VRRASEVTFAGRYPEMVANCISQDLSWSPAKQWEGILEE--VVGKGV-----ATAKK 519
Db 540 LQRAIKVGTPTAYEMVRNCMIQDLSWGPAPKNWENVLLSLGVAGGEGVGEELAPLAK 599
QY 520 EEIKVP 525
Db 600 ENVAAP 605

RESULT 9
AAR25474
ID AAR25474 standard; Protein: 502 AA.
XX
AC AAR25474;
XX
DT 15-JAN-1993 (first entry)
DE
DE Central fragment from potato GBSS.
KW Granule-bound starch synthase; amylopectin; amylose production;
KW inhibition.
XX
OS Solanum tuberosum.
XX
PN W09211376-A.
XX
PD 09-JUL-1992.
XX
PF 20-DEC-1991; 91WO-SE00892.
XX
PR 21-DEC-1990; 90SE-0004096.
XX
PA (AMYL-) AMYLOGENE HB.
XX
PI Hofvander P, Persson PT, Tallberg A, Wikstroem O;
XX
WPI; 1992-250096/30.
DR N-PSDB; AAQ26401.
XX
PT Modifying potatoes to form amylopectin starch - using an
XX anti-sense construct to inhibit granule-bound starch synthase
XX
PS Claim 1; Page 21; 46pp; English.
XX
CC A genomic library in EMBL3 was prepared using leaves of the potato
CC Binje. The library was screened with cDNA clones for the 5' and 3'
CC ends of the GBSS gene. A full-length clone of potato GBSS gene was
CC identified (wx311). A BglII-SpeI fragment (1"m") of the clone was
CC found to contain the central region of the gene and was cloned in
CC pUC13 to give pSm. Restriction of pSm with NsiI and HpaII gives
CC fragment II which was cloned in pURD184 to give pURDmitt. Further
CC restriction of pURDmitt with HpaI-SstI gives a 2549bp fragment
CC comprising exons and introns from the middle of the gene. The
CC fragment can be used to make antisense constructs to suppress amylose
CC formation in potatoes. See also Q26400-4.
XX
SQ Sequence 502 AA;

Query Match 39.5%; Score 1331.5; DB 13; Length 502;
Best Local Similarity 55.6%; Pred. No. 2.5e-96;
Matches 269; Conservative 55; Mismatches 125; Indels 35; Gaps 9;

QY 2 LDIWVAAVAPWSKTGGGLDVTGGLPTIELVKRHRVMTIAPRYDOYADAWDTSVVVDI- 60
Db 39 MNLIIFGVTVGPSTKGGGLPDVGLGLPALAARGHRVMTISPRYQYKDAWDTGYAVEVK 98
QY 61 MG---EKVRYFHSIKGVRHWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNKHKRFALFC 117
Db 99 VGDSTIEIVRFFHCYKRGVDRVDFVHPFLEKRVWGKTGSKTYGKAGLDYLDNLFSLFC 158
QY 118 KAAIEARVILPF-----GP-GEDEVFANDWHSALVPVLKDEYQPKGQFTKAKSVLA 169

```

```

Db 159 QAALEAPKVLNLNNSNYFSGPYGEDVLFANDWHTALPCYLKSMYQSRGIYLNKAVFC 218
QY 170 IHNIAFOGRMEERAFKDKLPQAAPDKLAFSDGAKVYVTEATPMBEDEKPLPLTGKTYKKI 229
Db 219 IHNIAFOGRFSFDEPLNLNLFDFRGSFDFIDGY-----PVKG---RKI 259
QY 230 NWLKGIIAADKLVTVSPNYATEIAAAGGVVELDTVIRAKGIEGIVNGMDIEENPKTD 289
Db 260 NWMKAGILESHRVTVSPYAGELVSPVHKGVELDSVLKTKCITGIVNGMDTQENPATD 319
QY 290 KFLSAPYD-QNSVYAGKAAKEALQAEGLPVDPTAPLFAFATIGRLEEGKGVDIILALPK 348
Db 320 KYTDVKYDITTVMDAKPPLLKALQAAVGLPDKKIPITGFTGRLEEGKGSILVAAIHK 379
QY 349 ILATPKVOIATILGTGKAAYEKLVAIGTKYKRAKGVKFSAPLAHMLTAGADEMLVPSR 408
Db 380 FIGL-DVQIVLVLTGCKEKEFEQIEQLEVLYPNKAGVAKFNVP LAHMITAGADEMLVPSR 438
QY 409 FEPCGLIQLOHAMRYGTVPVWASTGGLVDITVKEGVTGFHMGALNP--DKLDEADADALAAT 466
Db 439 FEPCGLIQLOHAMRYGTVPITCASTGGLVDITVKEGVTGFHMGAFNVECDVVDPADVLKIVTT 498
QY 457 VRR 470
Db 499 VARA 502

RESULT 10
AAB31175
ID AAB31175 standard; Protein: 238 AA.
XX
AC AAB31175;
XX
DT 02-APR-2001 (first entry)
DE
DE A granule bound starch synthetase II (GBSSII).
XX
KW Granule bound starch synthetase II; GBSSII; starch grain;
KW adenosine diphosphate glucose-alpha1,4-giucan alpha4-glucosyltransferase;
KW starch synthetase.
XX
OS Chlamydomonas reinhardtii.
XX
PN FR2793806-A1.
XX
PD 24-NOV-2000.
XX
PF 21-MAY-1999; 99FR-0006494.
XX
PR 21-MAY-1999; 99FR-0006494.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Dhulst C, Ball S;
XX
WPI; 2001-052291/07.
DR N-PSDB; AAC86955.
XX
PT New recombinant nucleic acid encoding fusion of starch synthase and
PT second component, useful in pharmaceutical and food compositions, is
PT targeted to starch granules -
XX
PS Claim 9; Fig 1; 52pp; French.
XX
CC The present sequence represents a granule bound starch synthetase II
CC (GBSSII). The specification describes a recombinant nucleic acid,
CC comprising a polynucleotide encoding an adenosine diphosphate
CC glucose-alpha1,4-giucan alpha4-glucosyltransferase or starch
CC synthetase, placed upstream of a sequence that encodes a polypeptide
CC of interest. The adenosine diphosphate glucose-alpha1,4-giucan
CC alpha4-glucosyltransferase protein can migrate to sites of
CC biosynthesis of starch grains in plant cells, becoming associated

```

with these grains. The recombinant nucleic acid sequence is used to target polypeptides of interest to starch grains.

Query Match	Best Local Similarity	Score 1194;	DB 22;	Length 238;
Matches 234;	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0;
QY	416	QLHAMHGYTPVYASTGGGLVDTKGVGTGFHMGALNPDKLDEADADALAAATVRRASVFA	475	
Db		:::		
QY	3	RVHAMHGYTPVYASTGGGLVDTKGVGTGFHMGALNPDKLDEADADALAAATVRRASVFA	62	
Db		:::		
QY	476	GGRYPMWANCISQDLSWSKPAQKBWGLLEEVYVYGGKVATAKKEIKVPVAEKIPGDLP	535	
Db		:::		
QY	63	GGRYPMWANCISQDLSWSKPAQKBWGLLEEVYVYGGKVATAKKEIKVPVAEKIPGDLP	122	
Db		:::		
QY	536	AVSYAPNTLKPVSASVEGNGAAAPKVGTTPAPMGANRATTPSGPSFAAAATPKVTTYKPAL	595	
Db		:::		
QY	123	AVSYAPNTLKPVSASVEGNGAAAPKVGTTPAPMGANRATTPSGPSFAAAATPKVTTYKPAL	182	
Db		:::		
QY	596	PATAKPKTAGLKLAGEASTTTSSENGAASNGNGNGASAKTSAAKPLVSAATRKSA	651	
Db		:::		
QY	183	PATAKPKTAGLKLAGEASTTTSSENGAASNGNGNGASAKTSAAKPLVSAATRKSA	238	
Db		:::		

RESULT 11	
AAB49307	
ID	AAB49307 standard; Protein; 534 AA.
XX	
XX	AAB49307;
XX	
XX	
DT	01-MAR-2001 (first entry)
XX	
DE	Wheat starch synthase GBSS protein.
XX	
DE	Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;
KW	food product; adhesive.
KW	
XX	
XX	
OS	Triticum aestivum.
XX	
XX	
PN	WO2000066745-A1.
XX	
XX	
PD	09-NOV-2000.
XX	
XX	
PF	28-APR-2000; 2000WO-AU00385.
XX	
XX	
PR	29-APR-1999; 99AU-0000052.

(CSIR) COMMONWEALTH SCI & IND RES ORG.
 (GOOD-) GOODMAN FIELDER LTD.
 (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
 XX
 Morell M, Li Z, Rahman S, Appels R;
 WPI; 2000-647602/62.
 Nucleic acid molecules encoding wheat starch synthase (WST)-I and
 WST-II, useful in modifying plant starch content and/or composition -
 Example 15; Fig 9; 21pp; English.
 The present invention relates to novel protein and coding sequences from
 wheat. The proteins are wheat starch synthases, designated SSII and
 SSIII. These can be used in the modification of plant starch content or
 composition, and to screen plants to identify mutations which affect
 starch content and composition. The starch can then be used in food
 products, such as flour, and in films, coatings, adhesives, building
 materials and packaging materials.

AA	Sequence	534 AA;
AA	Query Match	26.8%; Score 904; DB 21; Length 534;

[illegible]

RESULT 12
ABB92160
ID ABB92160 standard; Protein; 792 AA.

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1371.

KW Herbicidal; plant; agriculture; herbicide.

OS *Arabidopsis thaliana*.

XX
PN
WO200210210-A2.

XX PD 07-FEB-2002.

XX
PF 28-AUG-2001: 2001WO-EP09892

XX
PR 28-AUG-2001. 2001WA-FB008803

XX
PA (FARR \ HAYES AC

XX
PT
E
+
C
S
E
E

PI Tietjen K, Weidler M;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:47:01 ; Search time 16.8268 Seconds
(without alignments)
3719.278 Million cell updates/sec

Title: US-09-980-771a-3_COPY_58_708

Perfect score: 3370
Sequence: 1 ALDIYVVAEAVPWSKTGL.....SASKTSAAKPIVSAATRKA 651

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1481.5	44.0	608	S43341	starch synthase (E
2	1457.5	43.2	607	YUPOY	starch synthase (E
3	1415.5	42.0	603	S61504	glycogen(starch) s
4	1412.5	41.9	608	T10906	starch synthase (E
5	1398.5	41.5	609	JQ0703	glycogen(starch) s
6	1397.5	41.5	609	S11481	glycogen(starch) s
7	1395.5	41.4	603	YUBHY	glycogen(starch) s
8	1390	41.2	610	F86453	granule-bound star
9	1362	40.4	615	1 YWYTY	glycogen(starch) s
10	1355.5	40.2	608	T14731	glycogen(starch) s
11	1353.5	40.2	605	S07314	glycogen(starch) s
12	1194	35.4	238	T07921	glycogen(starch) s
13	902.5	26.8	641	T07668	probable starch sy
14	873.5	25.9	626	JQ2322	starch synthase (E
15	872	25.9	732	T01208	starch synthase (E
16	871	25.8	752	S61505	glycogen(starch) s
17	869	25.8	610	T06280	probable starch sy
18	861.5	25.6	622	T01414	starch synthase (E
19	847.5	25.1	788	T07667	glycogen(starch) s
20	839	24.9	491	T06798	probable starch sy
21	837	24.8	698	T01209	starch synthase (E
22	684.5	20.3	477	B95130	glycogen synthase
23	683.5	20.3	477	H98000	starch synthase (E
24	677.5	20.1	484	S40051	starch synthase (E
25	669	19.9	483	C86712	starch synthase (E
26	652	19.3	480	H98228	glycogen synthase
27	652	19.3	480	AF3057	glycogen synthase
28	651	19.3	486	H72321	glycogen synthase
29	637	18.9	477	A97176	glycogen synthase,

ALIGNMENTS

RESULT 1

S43341

starch synthase (EC 2.4.1.21) precursor - cassava

N;Alternate names: bacterial-glycogen synthase; glycogen synthase; starch synthase
C;Species: Manihot esculenta (cassava)

C;Date: 07-Sep-1994 #sequence_revision 03-Nov-1995 #text_change 16-Aug-2002

C;Accession: S43341

R;Salehuzzaman, S.N.I.M.; Jacobsen, E.; Visser, R.G.F.

Plant Mol. Biol. 23, 947-962, 1993

A;Title: Isolation and characterization of a cDNA encoding granule-bound starch synt

A;Reference number: S43341; MUID:94083565; PMID:8260633

A;Accession: S43341

A;Molecule type: mRNA

A;Residues: 1-608 <SAL>

A;Cross-references: EMBL:X74160; NID:g437041; PIDN:CAA52273.1; PID:g437042

C;Genetics:

A;Gene: GBSS; waxy

A;Genome: nuclear

C;Function:

A;Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producir

C;Superfamily: starch synthase

C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra

F;1-78/Domain: transit peptide (amyloplast) #status predicted <NP>

F;79-608/Product: ADPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 44.0%; Score 1481.5; DB 2; Length 608;

Best Local Similarity 55.0%; Pred. No. 4.4e-81;

Matches 299; Conservative 64; Mismatches 148; Indels 33; Gaps 9;

QY	2	LDIVMVAEAVPWSKTGLGVDVTGGLP	FIELVKRGHRVMTIAPYDQYADAWDTSVVVDI- 60
Db	82	MNLIFVGAEGPWSKTGLGVDVGLG	LPAMAARGHRVMTVSPRYDQYKDAWDTSVSVEIK 141
QY	61	MG---EKVYFHSIKKGVHRVWDHFW	FLAKVWGKTGSKLYGPRSGADYLDNHHKRALPC 117
Db	142	IGDRIETVRFHSYKRGVDVRFVDF	HPMFLKFWGKTGSKIYGRAGLDYQDNQLRSLIC 201
QY	118	KAATAEARVLPF-----GP-GE	DCVFVANDWHSALVPVLKDEYQPKGQFTKAKSVLA 169
Db	202	LAALAEAPRVNLNSSKNFSGPYG	EEVAFIANDWHTALLPCYLKAIYQPGIYKHAKVAPC 261
QY	170	INIIAQFGRMEEAFKDKLPQAAFD	KLAFSDYAKRVYTEATPMEEDEKPPLTGTYKKI 229
Db	262	INIIAQFGREAFSDPRLNLPDK	ESDFIDGYEK-----PVKG---RKI 304
QY	230	NWLKGITTAADKLIVTSNPYAT	EIAADAGGVELDTVIRAKGIEGIVNGMDIEWNPKT 289
Db	305	NMKAGILSDRLVTSVPYIAQ	EVISGVERGVELDNFIRKKTGIAGTINGMDYQWNPVD 364
QY	290	KFLSAPYDONSYYACGA	AKAALQAFGLGVPDPTAPLAFIFGLBEQKGVDIILALPKI 349
Db	365	KYIDHYDATVMDQAKPL	LKEALQAEVGLPDRNVPLTIGFIFGLBEQKGSDFVAAISO- 423

glycogen synthase
starch (bacterial
probable starch sy
starch synthase (E
glycogen synthase
starch synthase (E
glycogen synthase
starch synthase (E
starch synthase (E
soluble starch syn
hypothetical prote
glycogen synthase
glycogen synthase
glycogen (starch)
hypothetical prote


```

197 QAALEAPRVNLNSKIFYSGPYGEDVIFVANDWHSALIPCYLKSMYKSRGLYKNAKVAFC 256
170 IHNTAFCQRMWEEAFKD---FKLPQAAFPDKLAFSDGYAKVYTEATPMEDEKPLTGKTY 226
257 IHNIAYOGR---NAFSDFSILNLPDEFSSDFIDGYNK-----PCBG--- 296
227 KKIWLKGGIITAADKLTVSPNYAETIAADAAGGVVELDTVIRAKGIEGIVGMDIIEWNP 286
297 KKIWNKAGILESQVFVSPHYAKELISGEDRGVVELDNIIRSTGIGIVGMDNREWSP 356
287 KTDRLFLSAPYDONSFYACKAAAKALQALQELGVPDPTAPLPAFVIGRLBEEQKGVDIILAA 346
357 QTRGYIDVHNETTIVTEAKPLKLTQALQELGVPDSSIPFLGFIGRLBEEQKGSILVEAI 416
347 PKILATPKVQTAIILGTGKAAEKLVLNAIGTYKGRAGVVKFSPALAHMLTAGADFMVLP 406
417 AK-FADENVQIVLVLTGKKIMEKQIEVLEEKYPKAGITGTFNPSPLAHKIITAGADFIVP 475
407 SRPEPCGLIQLHAMHYCTVPVWASTGGIVDTVKEGVTGFHMGALNP--DKLDEADADALA 464
476 SRPEPCGLVQLHAMPYCTVPVIVSTGGIVDTVKEGYTGFGHAGPFDEVEDVDPDDVKLA 535
465 ATVRRASEVPAGGRYPMPWANCITSQDLSWSKPAQKWEGCLEEV-----VYKGG--VATA 517
536 ATVKRAIKTYGTQAMQIILNCMAQNPSSWKPAKLWEKALNLEVTGNVAGIDGDEIAPL 595
518 KKEEIKVP 525
596 AKENVATP 603

```

```

RESULT 4
Ti0906
starch synthase (EC 2.4.1.21) - sweet potato
N:Alternate names: starch synthase
C:Species: Ipomoea batatas (sweet potato)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C:Accession: F00906

```

```

Db      262  IHNIAYQGRRFAFSDFSLNLNLPDEYKGSFDFIDGYDK-----PVKG---RKI   304
QY      230  NWLKGIIAADKLVTWSPNYATEIADRAAGGVELDTVIRAKGIEGIVNGMDTEWNPXTD  289
Db      305  NWMKAGIREADVFTVSPNYAKELVSCVSGVELDNHINDCGITGICGMDTQEWNPATD  364
QY      290  KFLSPAYDONSYYAGKAAAKAALQAEGLGVLDPDPTAPLFAFVIGRLEEQRGVDIILAALPKI  349
Db      365  KYLAVKYDITTVMQAKPULKALQAAVGLPVDNRNPLIGFTGRLEEQRGSILLYAASKF  424
QY      350  LATPKVQIAILGTGKAAAYEKLVLNAGTYKYGKRAKVVWFSPAPLAHMLTAGADFMVPSRF  409
Db      425  ISM-DVQILILGTGKKKEQQIEQLVEMYDPDKARGVAKFNVLAHMITAGADFMVPSRF  483
QY      410  EPCGLIQLHAMHYGTVPVYVASTGGVLVDVTKEGVTFGHMGALNPD--KLDEADADALAATV  467
Db      484  EPCGLIQLHAMRYGTPCICASTGGVLVDVTKEGVTFGHMGAFNVDCETVPDPEVLKVITTV  543
QY      468  RRASEVFAAGGYRPEWVANCISQDISWSKPAOKWSGLL-----EEVYVGKGVGATAKK  519
Db      544  GRALAYGTLAFTEMIKNMSQELSWKGPAPKNNWETVLLSLGVAGSGPGEVE-BIAPLAK  602
QY      520  BEIKVP 525
Db      603  ENVATP 608

RESULT 5
QJ00703
glycogen(starch) synthase (EC 2.4.1.11) - rice
N:Alternate names: granule-bound starch synthase
C:Species: Oryza sativa (rice)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Aug-2002
C:Accession: JQ00703
R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
Nucleic Acids Res. 18, 5898, 1990
A:Title: Nucleotide sequence of rice waxy gene.
A:Reference number: JQ00703; MUID:91016948; PMID:2216792
A:Accession: JQ00703
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-609 <WAN>
A:Cross-references: EMBL:X53694
A:Experimental source: strain subsp. japonica Hengfeng
C:Genetics:
A:Gene: waxy
A:Introns: 114/1; 141/1; 174/1; 204/1; 225/2; 259/1; 295/3; 377/1; 436/1; 500/1; 529/1
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
C:Superfamily: starch synthase

```

[illegible]

[illegible]

RESULT 6

S11481
glycogen(starch) synthase (EC 2.4.1.11) precursor - rice
N:Alternate names: starch synthase; waxy protein
C:Species: *Oryza sativa* (rice)
C:Date: 05-Mar-1995 #sequence,revision 10-Nov-1995 #text_change 16-Aug-2002
C:Accession: S11481; S22519; #sequence,revision 10-Nov-1995 #text_change 16-Aug-2002
R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.
Nucleic Acids Res. 18, 5898, 1990
A:Title: Nucleotide sequence of rice waxy gene.
A:Reference number: JQ0703; MUID:91016948; PMID:2216792
A:Accession: S11481
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-609 <WAN>
A:Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA37732.1; PID:g577599
R:Okagaki, R.J.
Plant Mol. Biol. 19, 513-516, 1992
A:Title: Nucleotide sequence of a long cDNA from the rice waxy gene.
A:Reference number: S22519; MUID:92322986; PMID:1377969
A:Accession: S22519
A:Molecule type: mRNA
A:Residues: 1-609 <OKA>
A:Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403
R:Hirano, H.Y.; Sano, Y.
Plant Cell Physiol. 32, 989-997, 1991
A:Title: Molecular characterization of the waxy locus of rice (*Oryza sativa*).

Query Match	41.58;	Score 1397.5;	DB 2;	Length 609;
Best local Similarity	51.48;	Pred. No. 4.7e-76;		
Matches 281;	Conservative	Mismatches 147;	Indels 43;	Gaps 9;
QY	2	LDIVMAAEVAPWSK	TGGLG	DVTGGLPELVKGRHYMTIAPRYDOYADAWDTSVVVDI- 60
Db	83	MNVVFGAEMAPWSK	TGGLG	DVLGGLPPAMAAHGRVWVISPRYDQKDAWDTSVAAEIK 142
QY	61	---MGEKVRVYFHS	TKKGVHRVWIDHP	WFLAKVWGKTKSKLYGPRSGADYLDNHRKFALFC 117
QY	143	VADRYERVREFFHCY	KRGVDVRFIDHP	SFLERKVGKTEKIGYPTDGYDKDNQMRFSLLC 202
QY	118	KAATEAARVL----	PTGCP---GDCYFVAND	WHSALVPVLLKDEYOPKGOFTRAKSVLA 169
Db	203	QAALAPRILNLN	NNPYFKGTG	YEDVVFVCMWHITGPIASYLKNNYQFNGIYRNAKVAFC 262
QY	170	IHNIAFOGRMEEAFK	QUTKLPQAAFDK	LAFSDGYAKVYTEATPMEDEKPLLTGKTYKKI 229
Db	263	IHNISYQGRFAFEDY	PELNLSEFRSSFD	IDGY-----DTPVEG-----RKI 305
QY	230	NWLKGGIITAADKL	VTVSPNYATEIAAD	AAGVELDVTIRAKGIEGIYNGMDIEBNPKTD 289
Db	306	NWMKAGILEADRVLT	VSPPYAAEELISGI	ARGCELDINMLRTGITGVNGMVDSEWDPSK 365
QY	290	KFLSAPYDONS	VYAKAAKALQAE	LGLPVPDTAPFAFTGRLEEQKGVDIILAAALPKI 349
Db	366	KYITAKYDATTAE	AKALKALQAEAGL	PVDRKIPLTAFTIGRLEEQKGPDMVMAAIPEL 425
QY	350	LATPKVQIAILGT	KGAAYEKVLNATG	TKYKFAKGVVKFSAPLAHMLTAGADFMVLVPSRF 409
Db	426	M-QEDVQIVLLGT	GKKKFEKLLKSME	EYPGKVRVAVVKNAPLAHLIMAGADVLVAPSRF 484
QY	410	EPCGLIQLHAMHYG	IVPVVASTGG	LVDTVKEGVTGFHMGALNPD--KJDEADADALAATV 467
Db	485	EPCGLIQLQGMRYG	TPCCASTGG	LVDTVIEKGTGFHMGRLSVDCCKVPEPSDVKKVAATPL 544
QY	468	RRASEVFAGGRYP	EMVANCISODLS	WSKPAOKWEGILLEEVYVYKGGK-----VATAK 518
Db	545	KRAIKVTPAYTE	EMVRNCMNQDLS	WKGPAPKNWENVL--LGLGVAGSAPGIGDETAIPLA 602
QY	519	KPEIKVP 525		
Db	603	KENVAAP 609		

RESULT 7

YUBHY
glycogen(starch) synthase (EC 2.4.1.11) precursor - barley
N:Alternate names: starch synthase
C:Species: Hordeum vulgare (barley)
C:date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
C:Accession: S01727; S01728
R:Rohde, W.; Becker, D.; Salamini, F.
Nucleic Acids Res. 16, 7185-7186, 1988
A:Title: Structural analysis of the waxy locus from Hordeum vulgare.
A:Reference number: S01727; MUID:88303345; PMID:2970062
A:Accession: S01727
A:Molecule type: DNA
A:Residues: 1-603 <ROH>
A:Cross-references: EMBL:X07931; NID:g19126; PID:CAA30755.1; PID:g295809
A:Accession: S01728
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-603 <ROH2>
A:Cross-references: EMBL:X07932; NID:g19128; PID:CAA30756.1; PID:g19129
C:Genetics:
A:Gene: waxy
A:Introns: 106/3; 133/3; 166/3; 218/1; 251/3; 369/3; 429/3; 493/3; 522/3; 565/3
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose p
A:Pathway: starch biosynthesis
C:Superfamily: starch synthase
C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hex

F:1-72/Domain: transit peptide (amyloplast) #status predicted <TNP>
 F:73-603/Product: UDPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 41.4%; Score 1395.5; DB 1; Length 603;
 Best Local Similarity 53.3%; Pred. No. 6.1e-76;
 Matches 277; Conservative 71; Mismatches 141; Indels 31; Gaps 7;

```

QY 2 LDIVMAAEVAPWSKTGGGLDVTGGPLIELVKRGHRVMTIAPRYDOYADAWDTSVVVDIM 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 MNLVFGAEMAPWSKTGGGLDVLGGPLPMAANGHRVMSPRYDQTKDAMDTSVISEIK 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 ----GEKRYFHSIKKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYLDNKHKRALFC 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 VADEYERVRFFHCYKRGVDRVFDHPWLEKVGKKEIKYIGPDAGTDYEDNQORSLLC 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 KAAIEAARVL----P-CP-GEDCVFVANDHSAALVPVLLKDEYOPKQFTKAKSVLA 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 QAALEAPRIILNANNFYFSGPYGEDVVFVNCNDWHTGLLACYLKSNGIYRTAKVAF 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 IHNIAFGQRMWEAFKDTKLPOAAFDKLAFCGAKYVTEATPMEDEKPLTKYTKKI 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 IHNISYGRFSFDFAQLNLPDRFKSFDIDYDK-----PVEG---RKI 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 NMLKGIIAADKLVTVSPNYATEIAAAGGVDELTVIRAKGIEGIVNGMDIEWNP 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 NMKAGILOADKLVTVSPNYATEIAAAGGVDELTVIRAKGIEGIVNGMDIEWNP 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 KFLSAPYDQNSVYAGKAAKALQAEGLPVDPTAPFAFVIGRLEOKGVDVILALPKI 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 KFLAVNYDITTALEAKALKEALQAEGLPVDPTAPFAFVIGRLEOKGVDVILALPKI 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 LATPKVQIATLGTGKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRF 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 LKEEDVQIILLGTGKPKKLLSMEKPPKVRVVRNAPLAHOMAGADLLAVTSRF 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDVTKGVTGFHMGALNPD--KLDEADADALAATV 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 EPCGLIQLGMRGYPGVCASGTGGLVDVTKGVTGFHMGRLSDVNCNVPEADVKVATTL 538
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 468 RRASEVAGGRYPPEMVVANCISDLSKSPAKQWEGLLREV 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 539 KRRAVVGTPAQEMVKNQMIQDLSKSGPARKNEDVILLEL 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8

F86453
 granule-bound starch synthase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86453
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dwyer, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maili, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Status: preliminary
 A:Accession: F86453
 A:Molecule type: DNA
 A:Residues: 1-610 <S>
 A:Cross-references: GB:AE005172; NID:G6910568; PIDN:AAF31273.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: starch synthase

Query Match 41.2%; Score 1390; DB 2; Length 610;
 Best Local Similarity 52.7%; Pred. No. 1.3e-75;
 Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;

```

QY 2 LDIVMAAEVAPWSKTGGGLDVTGGPLIELVKRGHRVMTIAPRYDOYADAWDTSVVVDI- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 MSVIFGAEGVPSKTGGGLDVLGGPLPAAARHVRMTICPRVDYQKDAWDTCVVQIK 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 MGEK---VRYFHSIKKGVHRVWIDHPWFLAKVWGKSGKLYGPRSGADYLDNKHKRALFC 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 VGDKVENRFFHCYKRGVDRVFDHPFLAKVWGKSGKLYGPRSGADYLDNKHKRALFC 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 KAAIEAARVLPF-----GP-GEDCVFVANDHSAALVPVLLKDEYOPKQFTKAKSVLA 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 QAALEAPQVNLNLSKSYFSGPYGEDVVFVANDWHTALLPCYLKSNYQSRGYMNAKVYFC 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 IHNIAFGQRMWEAFKDTKL---P-CP-GEDCVFVANDHSAALVPVLLKDEYOPKQFTKAKSVLA 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 264 IHNIAFGRAFDDYLLNLPISFKSFD---FMDGYEK-----PVKG--- 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 227 KKNLKGIIAADKLVTVSPNYATEIAAAGGVDELTVIRAKGIEGIVNGMDIEWNP 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 304 RKNLWMAAILEAHRLVTVSPNYAQLISGVDGVELHKLKMKTVSGIINGMDVQWNP 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 KTKFELSAFYDQNSVYAGKAAKALQAEGLPVDPTAPFAFVIGRLEOKGVDVILAL 346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 STDYKIDIKDITVTDAPLKEALQAAVGLPVDVPGVIGFIRLEOKGVDVILAL 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 PKLATPKVQIATLGTGKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMV 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 SKEMGL-NVQMWILGTGKKEAQILEEFKFGKAVGAKFVNPVLAHMLTAGADFMV 482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 SRPEPGLIQLHAMHYGTVPVVASTGGLVDVTKGVTGFHMGALNP--DKLDEADADALA 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 483 SRPEPGLIQLHAMRYGTVPIVASTGGLVDVTKGVTGFHMGALNP--DKLDEADADALA 542
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 ATVRRASVYFAGGRYPPEMVVANCISDLSKSPAKQWEGLLREV 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 543 KAVTRAVYGTGSAQEMVKNQMIQDLSKSGPARKNEDVILLEL 600
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 525 PVAEK 529
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 601 PLAKE 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9

YUWY
 glycogen(starch) synthase (EC 2.4.1.11) precursor - wheat
 N:Alternate names: starch synthase
 C:Species: Triticum aestivum (common wheat)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 02-Aug-2002
 C:Accession: S16261; S33636
 R:Clark, J.R.; Robertson, M.; Ainsworth, C.C.
 Plant Mol. Biol. 16, 1099-1101, 1991
 A:Title: Nucleotide sequence of a wheat (Triticum aestivum L.) cDNA clone encoding th
 A:Reference number: S16261; MUID:9122506; PMID:1863765
 A:Accession: S16261
 A:Molecule type: mRNA
 A:Residues: 1-615 <CLA>
 A:Cross-references: EMBL:X57233; NID:921901; PIDN:CAA40509.1; PID:921902
 R:Ainsworth, C.; Clark, J.; Balesdon, J.
 Plant Mol. Biol. 22, 67-82, 1993
 A:Title: Expression, organisation and structure of the genes encoding the waxy protei
 A:Reference number: S33636; MUID:93271462; PMID:8499619
 A:Accession: S33636
 A:Molecule type: protein
 A:Residues: 71-75, 'X', 77-78 <AIN>
 C:Genetics:
 A:Gene: waxy
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose produc
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; glycogen/starch biosynthesis; glucosyltransferase; hexosyltra
 F:1-70/Domain: transit peptide (amyloplast) #status predicted <TNP>
 F:71-615/Product: UDPglucose-starch glucosyltransferase #status experimental <MAT>

[illegible]

RESULT 10
 T14731
 glycogen(starch) synthase (EC 2.4.1.11) precursor, granule-bound - sorghum
 N:Alternate names: starch synthase, granule-bound
 C:Species: Sorghum bicolor (sorghum)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
 C:Accession: T14731
 R:Hsing, Y.C.
 submitted to the EMBL Data Library, April 1995
 A:Reference number: Z18173
 A:Accession: T14731
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-608 <HS1>
 A:Cross-references: EMBL:U23945; NID:g1255713; PID:g1255714
 A:Experimental source: strain 12311; young seed
 C:Genetics:
 A:Genome: nuclear
 A:Note: Wx
 C:Function:
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing
 A:Pathway: starch biosynthesis
 C:Superfamily: starch synthase
 C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase
 F:1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>
 F:78-608/Product: UDPglucose-starch glycosyltransferase #status predicted <NAT>

QY	2	LDIVMVAEAVPWSKTTGGLGVGTGGPLIELVKGRHRYMTAPRYDQTDADAWDTSVVVDI- 60
DB	82	MNVVFGAEAPWSKTTGGLGVGLPFPMAAANGHRVMVSPRYDQYKDAMDTSVVSJGE 141
QY	61	MG----EKVRYEHSFKKGVRHWIDHPFLAKVWGCTGSKLYGPRSGADYLDNHRKRALFC 117
DB	142	MGDGYETVRPHCYKRGVDRVFDIDHPLFLERVWGKTEKIIYGPDAQTDYKDQNLRFSLJC 201
QY	118	KAATEAARVL-----PF--GP-GEDEVFVAMCHSALVPPVLLKDEYOPKQFTTKAKSVLA 169
DB	202	QAALAPRLLSNPNPYFSGPYGDEVVFCVNIWHTGPLSCYLKSNYQSNIGYDKDAKTAFC 261
QY	170	IHNIAFGRMWEEAFKTKLPQAAFDKLATSLGYAKYATTEATPMEDEKPLTKGTYKKI 229
DB	262	IHNISYQGRFAFSDFPELNLPEREKSSDFILGYEK-----PVEG---RKI 304
QY	230	NWLKGGIITADKLVTPSPNTATEITAAADAAGVELDTVIRAKGIEGIVNGMDEIENPKTD 289
DB	305	NWMKAGILEADRVLTSPYAEELISGLTARGCELDNIMRLTGITGVNGMDVSEWSPDKD 366
QY	290	KFLSAPPQDQNSVYAGKAAAKALQAEUGLQVDPDTAPLPAFAFGRLEEQKGVDIILAAALPKI 349
DB	365	KYIAVKYDVSTAVEAKALKALQAEVLGPDYRKIPVAFATIGRLEEQKGPDMMAAIP-L 423
QY	350	LATPKVQIATLGTCKAAEYKELVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMILVPSRF 409
DB	424	IMEEDIQILVLGTGKKFERMLMSAEKYPDKVRAVVFENALAHIMAGADLLAVTSRF 483
QY	410	EPCGLIQIOLHAMHYGTPVVPVASTGGVLDTVKEGVTGFHMGALNPD--KLDEADADALAAATV 467
DB	484	EPCGLIQIOLQGRYGTGTPCACASTGGVLDTIIEGKTGFHMGRLSVDCCNVVPEADVKVAVTL 543
QY	468	RRASEVFAGGRYPENVANCISODLSWSPKPAQKWEGILLEE--VYVKGG-----VATAKKE 520
DB	544	KRAIKVGTPTAYEEMVKCMIQDLSWKGPAKINWENVLISLGVAGGEPGEIEGEEIAPLAKE 603
QY	521	EIKVP 525
DB	604	NVAP 608
RESULT 11		
S07314		
glycogen(starch) synthase (EC 2.4.1.11) precursor - maize		
N:Alternate names: starch synthase; UDP-glucose starch glycosyltransferase, st		
C:Species: Zea mays (maize)		
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Aug-2002		
C:Accession: S07314		
R:Kloosgen, R.B.; Gierl, A.; Schwarz-Sommer, Z.; Saedler, H.		
Mol. Gen. Genet. 203, 237-244, 1986		
A:Title: Molecular analysis of the waxy locus of Zea mays.		
A:Reference number: S07314		
A:Accession: S07314		
A:Molecule type: DNA		
A:Residues: 1-605 <RLO>		
A:Cross-references: EMBL:X03935; NID:g22509; PIDN:CAR27574.1; PID:g1644339		
A:Experimental source: line C		
A:Note: translation of the nucleotide sequence is not complete		
A:Note: part of this sequence, including the amino end of the mature protein,		
C:Genetics:		
A:Gene: waxy		
A:Introns: 107/3; 134/3; 167/3; 197/3; 219/1; 252/3; 289/2; 370/3; 431/3; 495/		
C:Function:		
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose p		
C:Superfamily: starch synthase		
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltra		
F;1-72/Domain: transit peptide (amyloplast) #status predicted <TNP>		
F;73-605/Product: glycogen (starch) synthase #status experimental <MAI>		
Query Match 40.2%; Score 1353.5; DB 1; Length 605;		
Best Local Similarity 50.4%; Pred. No. 2e-73;		
Matches 275; Conservative 79; Mismatches 153; Indels 39; Gaps 11		